Die



Ausführliche Prospekte durch jede Buchhandlung oder direkt von der Verlagshandlung.

Verzeichnis der Bände der „Sammlung Schubert".

(Voraussichtlicher Erscheinungstermin in Klammer.)

Die Preise betragen für das gebund. Exempl. ca. M. 2.50 bis M. 5.00.

Band I: **Elementare Arithmetik und Algebra.** Von Prof. Dr. H. Schubert in Hamburg. (Juni 1899.)

„ II: **Elementare Planimetrie.** Von Prof. W. Pflieger in Münster. (Mai 1900.)

„ III: **Ebene und sphärische Trigonometrie.** Von Dr. F. Bohnert in Hamburg. (Oktbr. 1899.)

„ IV: **Konstruierende u. beschreibende Stereometrie.** Von Prof. Dr. G. Holzmüller in Hagen. (Juni 1899.)

„ V: **Niedere Analysis.** Von Prof. Dr. H. Schubert in Hamburg. (Mai 1900.)

„ VI: **Algebra, Determinanten und elementare Zahlentheorie.** Von Dr. O. Pund in Altona. (Juni 1899.)

„ VII: **Ebene Geometrie der Lage.** Von Dr. R. Böger in Hamburg. (Oktober 1899.)

„ VIII: **Analytische Geometrie der Ebene.** Von Prof. Dr. M. Simon in Strafsburg. (Oktober 1899.)

„ IX: **Analytische Geometrie des Raumes.** Von Prof. Dr. M. Simon in Strafsburg. (Mai 1900.)

„ X: **Differentialrechnung.** Von Prof. Dr. F. Meyer in Königsberg. (Mai 1900.)

„ XI: **Integralrechnung.** Von Prof. Dr. F. Meyer in Königsberg. (Mai 1900.)

Band XII: **Darstellende Geometrie.** Von Dr. J. Schröder in Hamburg. (Oktober 1899.)

„ XIII: **Differentialgleichungen.** Von Prof. Dr. L. Schlesinger in Klausenburg. (Januar 1900.)

„ XIV: **Praxis der Gleichungen.** Von Prof. C. Runge in Hannover. (Oktober 1900.)

„ XV: **Elemente der Astronomie.** Von Dr. E. Hartwig in Bamberg. (Mai 1900.)

„ XVI: **Mathematische Geographie.** Von Dr. E. Hartwig in Bamberg. (Mai 1900.)

„ XVII: **Berechnende Stereometrie.** Von Prof. Dr. G. Holzmüller in Hagen. (Januar 1900.)

„ XVIII: **Geschichte der Mathematik.** Von Prof. Dr. R. Haufsner in Giefsen. (Oktober 1899.)

„ XX: **Versicherungsmathematik.** Von Dr. F. Paul in Budapest. (Mai 1900.)

Die Sammlung wird fortgesetzt.

Sammlung Schubert VI

Algebra

mit Einschluss

der elementaren Zahlentheorie

Von

Dr. Otto Pund

Oberlehrer in Altona

Leipzig

G. J. Göschensche Verlagshandlung

1899

Vorwort.

Die vorliegende Darstellung der Algebra nimmt eine vermittelnde Stellung ein zwischen den Elementarlehrbüchern und den ausführlichen Handbüchern der höheren Algebra, von denen wir seit dem letzten Jahrzehnt auch hervorragende von deutschen Verfassern besitzen.*) Während sie die elementaren Teile der Wissenschaft, wie sie im ersten Bande dieser Sammlung aus der Hand des in der mathematischen Welt wohlbekannten Leiters zur Darstellung kommen, voraussetzt, sucht sie das Verständnis für die schwierigeren Gebiete anzubahnen und diese nach Gesichtspunkten zu behandeln, deren Wichtigkeit in neuerer Zeit immer mehr und mehr hervorgetreten ist, die aber weiteren Kreisen so gut wie unbekannt geblieben sind.

Vor allem kommt als solcher Gesichtspunkt die Gruppentheorie in Betracht, die, eine eigentliche Schöpfung unseres Jahrhunderts, ihren beherrschenden Einfluſs auf fast allen Gebieten der Mathematik, nicht etwa bloſs in den neu entstandenen Theorieen, sondern weit bis in die Anfangsgründe hinein, gezeigt hat, so daſs sie jetzt auch in den elementaren Lehrbüchern nicht mehr fehlen darf. Überall, wo es anging, habe ich deswegen die gruppentheoretischen Methoden in den Vordergrund treten lassen, so daſs der Leser, wenn er das Auftreten derselben formalen Schlüsse bei ganz verschiedenartigen Untersuchungen kennen gelernt hat, nun von selbst die Möglichkeit einsehen, vielleicht auch gar das Bedürfnis fühlen wird, diese Schlüsse von der Besonderheit der jeweiligen Gruppenelemente und

*) H. Weber, Lehrbuch der Algebra. In zwei Bänden. Braunschweig. Zweite Auflage 1898 — E. Netto, Vorlesungen über Algebra. Im Erscheinen begriffen. Leipzig. — Von älteren Werken ist der verdienstvolle Cours d'Algèbre supérieure von Serret hervorzuheben, der in deutscher Übersetzung von Wertheim unter dem Titel Handbuch der höheren Algebra (Leipzig, 1878|79) erschienen ist.

ihrer Zusammensetzung loszulösen, um so zu einer ganz abstrakten Theorie der Gruppen zu gelangen. Eine solche selbst hat jedoch in diesem Buche keinen Platz mehr gefunden; es mufs daher für ein eingehenderes Studium auf die vorzügliche Darstellung im zweiten Bande von H. Weber, Lehrbuch der Algebra, verwiesen werden.

Zeigt sich der Gruppenbegriff in fast allen Zweigen der Mathematik von Bedeutung, so ist der Begriff der Modulsysteme ein speziell algebraischer und zahlentheoretischer. Obwohl er auf einem etwas entlegenen Gebiet wie dem der algebraischen Zahlen durch die Arbeiten von Dedekind und Kronecker erwachsen ist, so tritt er in einfachster Form doch schon in der elementaren Zahlentheorie auf, und deswegen wird es kein Bedenken erregen, wenn hier der Versuch gemacht wird, den Begriff methodisch einzuführen. Die Behandlungsweise der linearen Kongruenzen wird die Überlegenheit des Verfahrens gegenüber den sonstigen Lösungsmethoden zur Anschauung bringen. Auch bei der Behandlung der linearen Gleichungssysteme bietet sich dann ein einfacher Weg zur Entwickelung des Determinantenbegriffes dar, zu dem man sonst auf völlig unvermittelte Art gelangt. Es ist hierbei besonderes Gewicht darauf gelegt worden, dafs nicht zuerst zu einer gekünstelten Begriffsbildung Zuflucht genommen wird, deren Wichtigkeit sich erst nachträglich mehr durch Verifikationen als natürliche Beweise herausstellt, sondern dafs statt dessen die Kette notwendiger Schlüsse vorausgeschickt wird. Die hier gegebene analytische Darstellung der Determinante auf Grund ihrer Eigenschaften rührt von Kronecker her.

Was die Auswahl des Stoffes betrifft, so habe ich nicht das Bestreben gehabt, in aller und jeder Hinsicht vollständig zu sein. Dazu ist der Stoff zu umfangreich; auch kann in Bezug auf manche Untersuchungen, wie die orthogonale Transformation quadratischer Formen, auf andere Bände dieser Sammlung*) verwiesen werden.

Altona, im Juli 1899.

Der Verfasser.

---

*) Z. B. auf Bd. VIII, IX. Simon, Analytische Geometrie der Ebene und des Raumes.

Inhaltsverzeichnis.

I. Abschnitt.

Rationale Funktionen.

§ 1. Rationale Rechenoperationen und Ausdrücke.

Unter einer rationalen Rechenoperation versteht man ein Rechenverfahren, das sich aus folgenden einfachen oder elementaren Rechenarten zusammensetzen läfst: der Addition und Subtraktion, der Multiplikation und der Division, die man schon von Alters her als die Grundrechnungsarten oder die vier Species bezeichnet.*) Hierbei ist zu bemerken, dafs der Subtraktion in der Algebra eine selbständige Stellung nicht zukommt, weil der Begriff der Addition bei der Einführung der negativen Zahlen so allgemein gefafst werden kann, dafs er die Subtraktion in sich begreift, und dafs auch das Potenzieren mit einer ganzen positiven oder negativen Zahl nicht als besondere Operation erwähnt zu werden braucht, weil es mit einer mehrmaligen Anwendung der Multiplikation oder Division gleichbedeutend ist. Handelt es sich nur um Operationen mit ganzen Zahlen, so ist auch die Multiplikation nicht als eine einfache Operation anzusehen, weil sie einer mehrfachen Addition gleichkommt. Dies gilt aber nicht mehr für den erweiterten Zahlbegriff, wie er in der Algebra im Gegensatz zur Zahlentheorie in Betracht kommt.

Mit Hülfe der rationalen Rechnungsarten ist es möglich, aus der Einheit alle positiven und negativen Brüche herzustellen, deren Gesamtheit man auch als rationale Zahlen

*) Vgl. Bd. I dieser Samml., Abschn. II. Anm. des Leiters.

bezeichnet. Sind mit einer Anzahl bestimmter Zahlen eine Anzahl von rationalen Operationen auszuführen, so ergiebt sich als Resultat wieder eine ganz bestimmte Zahl, nur den einen Fall ausgenommen, dafs eine Division durch Null hierbei vorkommt.

Wenn jedoch die Zahlen unbestimmt gelassen und demgemäfs durch Buchstaben bezeichnet werden, so kann es sich nur darum handeln, den so definierten rationalen Ausdruck durch Anwendung der arithmetischen Gesetze auf eine möglichst einfache Form zu bringen. Über diese Art der Reduktion läfst sich, wie in der allgemeinen Arithmetik gelehrt wird,*) hier aber doch noch kurz auseinandergesetzt werden soll, folgendes feststellen:

I. Zunächst kann der rationale Ausdruck, wenn die Operation der Division bei seiner Berechnung überhaupt in Anwendung kommt, so umgeformt werden, dafs diese nur einmal und zwar zuletzt ausgeführt zu werden braucht. Diese Möglichkeit beruht darauf, dafs sich algebraische Summen, deren Glieder, und Produkte, deren Faktoren alle oder teilweise Quotienten sind, stets in Quotienten verwandeln lassen. Nennt man einen rationalen Ausdruck, bei dessen Berechnung nur die Operationen der Summation und Multiplikation erforderlich sind, einen ganzen rationalen Ausdruck, einen solchen, bei dem auch Divisionen in Betracht kommen, einen gebrochenen, so kann man den Satz aussprechen: Jeder gebrochene rationale Ausdruck läfst sich als Quotient von zwei ganzen rationalen Ausdrücken darstellen. Über diese letzteren gilt nun der folgende Satz:

II. Jeder ganze rationale Ausdruck kann so umgeformt werden, dafs die Multiplikationen zuerst und die Summationen zuletzt ausgeführt werden. Dies ergiebt sich aus dem arithmetischen Satze, dafs sich Produkte aus algebraischen Summen stets als algebraische Summen darstellen lassen.

Werden rationale Ausdrücke wieder rationalen Rechenoperationen unterworfen, so entstehen stets wieder rationale Ausdrücke; die Division durch Null mufs hierbei jedoch ausgeschlossen werden.

---

*) Vgl. Bd. I dieser Samml., § 16. (A. d. L.).

§ 2. Rationale Funktionen einer Veränderlichen.

Kommt in einem rationalen Ausdruck nur eine unbestimmte Gröſse x vor, so nennt man diesen eine rationale Funktion der Variabeln oder Veränderlichen x. Eine solche kann also gebildet werden, indem man mit der Unbestimmten x und den rationalen Zahlen rationale Operationen vornimmt. Da aber die rationalen Zahlen alle aus der Einheit hervorgehen, so kann man eine rationale Funktion auch ansehen als einen rationalen Ausdruck, der aus dieser Einheit und der Grösse x durch die Grundrechnungsarten gebildet werden kann. Je nachdem nun bei der Bildung die Operation der Division in Anwendung kommt oder nicht, nennt man die Funktionen ebenso wie die rationalen Ausdrücke gebrochene oder ganze. Die ersteren lassen sich stets als Quotienten zweier ganzen Funktionen darstellen. Um die allgemeine Form dieser ganzen Funktionen zu ermitteln, beachten wir, daſs zunächst Multiplikationen und darauf Summationen in Betracht zu ziehen sind. Da die Multiplikationen nur mit der Unbestimmten x auszuführen sind, weil die Einheit keine Veränderung hervorbringt, so gelangt man durch ihre Anwendung zu lauter Potenzen von x mit ganzen positiven Exponenten. Bringt man nun die Summation zur Ausführung, so ergiebt sich aus der Einheit eine ganze positive oder negative Zahl, während mehrfach vorkommende gleiche Potenzen von x zu einem einzigen Gliede oder Term vereinigt werden können, das also die Form $a x^{\nu}$ hat, wo a eine ganze positive oder negative Zahl bedeutet und der Koeffizient der Potenz x^{ν} genannt wird, während ν nur eine positive ganze Zahl sein kann. Es ist zweckmäſsig, auch das von x freie Glied, das aus der Summation von Einheiten entsteht, in dem wir es einer bekannten Konvention zufolge, als den Koeffizienten der nullten Potenz von x ansehen, in diese Form der Glieder mit einzubegreifen, und also für ν auch den Wert o zuzulassen. Ein Aggregat von lauter Gliedern, bei denen die Potenzexponenten verschieden sind, ist nun, so lange x als unbestimmte Gröſse angesehen wird, keiner weiteren Vereinfachung mehr fähig. Wohl aber kann man die Glieder in bestimmter Weise nach der Gröſse der Potenzexponenten anordnen, entweder nach fallender oder nach steigender. Der

gröfste Potenzexponent n spielt bei vielen Untersuchungen eine wichtige Rolle, und man schreibt nach ihm der ganzen Funktion den Grad oder die Ordnung n zu; dabei wird dann die Annahme gemacht, dafs der Koeffizient dieser höchsten nten Potenz nicht verschwindet. Eine Funktion nten Grades hat hiernach, wenn wir mit a_0 den Koeffizienten von x^n, a_1 den von x^{n-1}, allgemein a_k den von x^{n-k} wo k die Werte 0,1..n anehmen kann, endlich mit a_n das von x freie Glied bezeichnen, die Form:

$$F(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_k x^{n-k} + \ldots + a_{n-1} x + a_n$$

wo $a_0 \neq 0$ ist, während die übrigen Koeffiizenten alle oder teilweise verschwinden können.

Der niedrigste Grad, den eine ganze Funktion besitzen kann, ist der erste; eine solche Funktion, die auf die Form $ax + b$ gebracht werden kann, wo a und b die Koeffizienten bedeuten, wird auch eine ganze lineare Funktion genannt. Ganze Funktionen zweiten Grades nennt man häufig quadratische, dritten Grades auch kubische, vierten Grades biquadratische. Manchmal ist es auch zweckmäfsig, von einer ganzen Funktion nullten Grades zu sprechen, worunter dann nur ein Koeffizient verstanden wird, der überhaupt mit keiner Unbestimmten verbunden ist. Jede gebrochene rationale Funktion von x läfst sich in der Form $\frac{f(x)}{g(x)}$ darstellen, wo f(x) und g(x) ganze Funktionen von x sind, mit denen man, wie wir später erkennen werden, unter Umständen noch gewisse Vereinfachungen vornehmen kann (s. § 56). Sind Zähler und Nenner vom ersten Grade, läfst sich also die Funktion darstellen durch $\frac{ax + b}{cx + d}$ mit den Koeffizienten a, b, c, d, so nennt man sie eine gebrochene lineare Funktion.

---

§ 3. Rationale Funktionen mehrerer Variablen.

Der Begriff der rationalen Funktion, den wir soeben für eine einzige Variable x kennen gelernt haben, läfst sich auf eine beliebige Anzahl n von Veränderlichen ausdehnen,

die wir mit $x_1, x_2, \ldots x_n$ bezeichnen wollen. Unter einer rationalen Funktion der Unbestimmten $x_1, x_2, \ldots x_n$ verstehen wir jeden Ausdruck, der aus ihnen und der Einheit durch Anwendung der rationalen Rechnungsarten hervorgeht. Eine ganze Funktion ist eine solche, bei deren Bildung nur die Addition und Multiplikation in Betracht kommen, während bei einer gebrochenen auch mindestens eine Division auszuführen ist. Eine solche läfst sich stets als Quotient von zwei ganzen Funktionen darstellen.

Um die allgemeine Form der ganzen Funktionen von $x_1, x_2, \ldots x_n$ zu finden, haben wir zunächst nur die Multiplikation in Betracht zu ziehen. Da hierbei aber nicht nur die einzelnen Veränderlichen mit sich selbst, sondern auch mit einander zu verknüpfen sind, so ergiebt sich als allgemeine Form das Potenzprodukt $x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$, wo $\nu_1, \nu_2, \ldots \nu_n$ ganze, nicht negative Zahlen sind, von denen nur alle nicht zugleich verschwinden können. Da aber aus der Einheit durch Addition ein ganzzahliger positiver oder negativer Koeffizient entsteht, so kann man diesen als Koeffizienten eines Potenzproduktes mit lauter verschwindenden Potenzexponenten ansehen. Wird nun auf alle Potenzprodukte noch eine Summation ausgeführt, so können alle, die dieselben Veränderlichen und jede mit demselben Exponenten besitzen, zu einem Gliede von der Form $a x_1^{\nu_1} x_2^{\nu_2} \ldots x_n^{\nu_n}$ zusammengefafst werden, wobei der Koeffizient a eine positive oder negative ganze Zahl bedeutet. Der gröfste Wert, den die Summe der Potenzexponenten erreicht, wird der Grad oder die Ordnung der ganzen Funktion genannt. Der niedrigste Grad, der vorkommen kann, ist der erste. Die Funktion kann in diesem Falle auf die Form gebracht werden:

$$a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$$

mit positiven oder negativen Koeffizienten $a_0, a_1, \ldots a_n$ und wird als eine lineare bezeichnet. Funktionen zweiten, dritten und vierten Grades nennt man auch quadratische, kubische und biquadratische. Haben alle Glieder einer ganzen Funktion dieselbe Ordnung, so nennt man diese eine homogene Funktion oder eine Form und unterscheidet nach der Ordnung lineare, quadratische, kubische, biquadratische und solche höheren Grades. Eine lineare Form unterscheidet sich von der oben hingeschriebenen

linearen Funktion nur durch das Fehlen des Koeffizienten a_0. Ist $F(x_1, x_2, .. x_n)$ eine Form n ter Ordnung, so gilt die Gleichung

$$F(tx_1, tx_2, .. tx_n) = t^n F(x_1, x_2, .. x_n).$$

Jede nicht homogene ganze Funktion läfst sich als eine Summe von Formen darstellen, die man dann nach der Höhe ihrer Ordnung aufeinander folgen lassen kann.

Bei einer solchen Anordnung bleibt die Stellung der Glieder in jeder Form immer noch willkürlich. Eine vollkommen bestimmte Anordnung der Glieder, nicht nur der homogenen, sondern überhaupt jeder ganzen Funktion allgemein, läfst sich auf folgende Weise erzielen. Wir setzen unter den Variabeln eine bestimmte Ordnung fest, so dafs eine beliebige x_1 den ersten, x_2 den zweiten u. s. w. schliefslich x_n den letzten n ten Platz erhält. Zunächst berücksichtigen wir nur die erste Veränderliche x_1 und ordnen sämtliche Glieder der Funktion so an, dafs diejenigen den Vortritt erhalten, die x_1 in einer höheren Potenz besitzen. Zur Ordnung derjenigen, die x_1 in gleicher Potenz oder auch garnicht enthalten, benutzen wir dann die zweite Variable x_2, und stellen die Glieder voran, die von ihr eine höhere Potenz haben. So fährt man weiter fort, bis schliefslich die letzte Variable x_n zur Anordnung der Glieder herangezogen wird, deren Stellung vorher zweifelhaft war, weil sie $x_1, x_2, .. x_{n-1}$ in denselben Potenzen enthielten. Dann ist aber die Anordnung eine völlig eindeutige, da mehrere Glieder mit gleichen Potenzexponenten in allen Veränderlichen nicht vorkommen können, sondern als in ein einziges Glied zusammengefafst vorausgesetzt werden. Der Gesichtspunkt, nach dem die Glieder geordnet sind, kann auch in folgender Weise ausgedrückt werden: Das Glied $a x_1^{\alpha_1} x_2^{\alpha_2} .. x_n^{\alpha_n}$ steht vor oder hinter dem Gliede $b x_1^{\beta_1} x_2^{\beta_2} .. x_n^{\beta_n}$, je nach dem die erste nicht verschwindende der Differenzen $\alpha_1 - \beta_1$, $\alpha_2 - \beta_2$, .. $\alpha_n - \beta_n$ einen positiven oder einen negativen Wert hat.

---

§ 4. Begriff der rationalen Funktion in erweiterter Bedeutung.

Häufig gebraucht man den Begriff der rationalen Funktionen in weiterer Bedeutung, als wir ihn den vorhergehenden Paragraphen entwickelt haben. Da es nämlich oft vorkommt, dafs einige der Variabeln $x_1, x_2, \ldots x_n$ eine ausgezeichnete Rolle bei gewissen Untersuchungen spielen, z. B. sich allein ändern, während andere $u_1, u_2, \ldots u_m$ ihren Wert beibehalten, so kann man die absolute rationale Funktion der $n+m$ Unbestimmten $x_1, x_2, \ldots x_n, u_1, u_2, \ldots u_m$ als relative rationale Funktion von $x_1, x_2, \ldots x_n$ ansehen, deren Koeffizienten dann absolute rationale Funktionen von $u_1, u_2, \ldots u_m$ sind. Der gewöhnlichste Fall ist der, dafs die Koeffizienten selbst diese unbestimmten Gröfsen sind. Die Gröfsen, aus denen sich die Koeffizienten rational ableiten lassen, nennt man oft ihren Rationalitätsbereich. Bei den rationalen Funktionen, wie wir sie in den vorigen Paragraphen betrachtet haben, ist dieser Rationalitätsbereich aus der Einheit gebildet, aus der sich ja durch rationale Operationen alle rationalen Zahlen ableiten lassen. Diesen Rationalitätsbereich nennt man auch den natürlichen. Eine ganze rationale Funktion, wie wir sie in den vorhergehenden beiden Paragraphen verstanden hatten, ist aber nicht mit einer ganzen Funktion im natürlichen Rationalitätsbereich identisch, da diese Funktion ja auch gebrochene rationale Koeffizienten haben kann. Um jene rationale Funktion mit ganzzahligen Koeffizienten zu kennzeichnen, werden wir sie fortan eine ganze ganzzahlige Funktion nennen. Eine rationale Funktion, bei deren Bildung mit gewissen Variabeln $x_1, x_2, \ldots x_n$ nur die Operationen der Addition und Multiplikation, mit den übrigen $u_1, u_2, \ldots u_m$ dagegen auch noch die Division in Betracht kommt, kann eine ganze Funktion von $x_1, x_2, \ldots x_n$ mit dem Rationalitätsbereich $(u_1, u_2, \ldots u_m)$ ihrer Koeffizienten genannt werden.

Wenn die Gröfsen $u_1, u_2, \ldots u_m$ des Rationalitätsbereiches nicht unabhängig von einander, sondern durch algebraische Beziehungen mit einander verknüpft sind, so nennt man den Rationalitätsbereich auch einen Gattungsbereich oder einen Körper (§ 99).

II. Abschnitt.

Arithmetische Eigenschaften rationaler Funktionen.

---

§ 5. Abkürzung mit Hilfe des Summen- und Produktzeichens.

Die Formeln der Algebra gewinnen nicht nur an Kürze, sondern auch an Übersichtlichkeit, wenn man sich gewisser Abkürzungen bedient, deren Gebrauch wir jetzt erläutern wollen. Da wir diese später fast durchweg anwenden, so empfehlen wir dem Leser, sich bald mit ihnen vertraut zu machen.

Sind eine Reihe von Gröfsen $a_1, a_2, \dots a_n$, deren Bezeichnung sich nur durch die Indices $1, 2, \dots n$ unterscheidet, zu summieren oder zu multiplizieren, so gebraucht man für die Summe

$$a_1 + a_2 + \dots + a_{n-1} + a_n$$

die Bezeichnungen

$$\sum_1^n{}_i\, a_i, \quad \sum_{i=1}^{i=n} a_i \quad \text{oder} \quad \sum_i a_i \qquad (i = 1, 2, \dots n)$$

für das Produkt

$$a_1\, a_2\, a_3 \dots a_n$$

ähnlich so die Abkürzungen:

$$\prod_1^n a_i, \quad \prod_{i=1}^{i=n} a_i \quad \text{oder} \quad \prod_i a_i \qquad (i = 1, 2, \dots n).$$

Dem Leser wird es nicht schwer fallen, die folgenden Formeln zu beweisen:

$$(1) \quad \sum_{(i)} (a_i + b_i) = \sum_{(i)} a_i + \sum_{(i)} b_i$$

$$(2) \quad \prod_{(i)} a_i b_i = \prod_{(i)} a_i \cdot \prod_{(i)} b_i$$

$$(3) \quad \sum_{(i)} m a_i = m \sum_{(i)} a_i$$

$$(4) \quad \prod_{(i)} a_i^m = \left(\prod_{(i)} a_i\right)^m$$

$$(5) \quad \prod_{(i)} a^{b_i} = a^{\sum_{(i)} b_i}$$

bei denen die Indices auf der rechten Seite genau dieselben Werte durchlaufen müssen wie auf der linken.

Die ganzen Funktionen

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n,$$
$$a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + a_n x^n$$

können kurz in der Form

$$\sum_i a_i x^{n-i} \quad \text{oder} \quad \sum_i a_i x^i \qquad (i = 0, 1, .. n),$$

das Produkt

$$(x + x_1)(x + x_2) \ldots (x + x_n)$$

durch

$$\prod_i (x + x_i) \qquad (i = 1, 2, .. n)$$

dargestellt werden. Die Gleichungen

$$(6) \quad \sum_i i = \frac{n(n+1)}{2}, \quad \sum_i a^i = \frac{a^{n+1} - 1}{a - 1} \qquad (i = 0, 1, .. n)$$

stellen die Hauptformeln der Theorie der arithmetischen und geometrischen Reihen dar, wie sie in der niederen Analysis abgeleitet werden (s. Bd. V dieser Samml.). Der Binomialsatz wird ausgedrückt durch

$$(7) \quad (a + b)^n = \sum_i n_i a^{n-i} b^i = \sum_i n_i a^i b^{n-i} \quad (i = 0, 1, .. n).$$

Die Binomialkoeffizienten sind definiert durch

$$(8)\quad n_0 = n_n = 1,\; n_i = \frac{n(n-1)\ldots(n-i+1)}{1.2\;\ldots\;i} = n_{n-i}$$
$$(i = 1, 2, \ldots n)$$

und wir geben hier ohne Beweis einige für sie geltende Relationen, indem wir auf Bd. V der Sammlung verweisen.

$$(9)\quad (i+1)_{k+1} = \sum_h h_k \qquad (h = k, k+1, \ldots i-1, i.)$$

$$(10)\quad \sum_h (-1)^h i_h h_k = 0 \quad (h = k, k+1, \ldots i-1, i;\; i > k).$$

Ist $i = k$, so hat die linke Seite nur ein einziges Glied und ist also gleich $(-1)^i$.

Eine mehr symmetrische Form des Binomialsatzes ist die folgende:

$$(11)\qquad (a+b)^n = \sum_{i,k} \frac{\Pi(n)}{\Pi(i)\,\Pi(k)}\, a^i\, b^k$$
$$(i, k = 0, 1 \ldots n;\; i+k = n),$$

wo trotz der beiden Indices i und k nur eine einfache Summe vorliegt, weil $i + k = n$ ist. Hierbei ist gesetzt

$$(12)\quad \Pi(o) = 1, \quad \Pi(n) = 1.\,2 \ldots n.$$

In dieser Form läſst sich die Formel zum Polynomialsatz erweitern:

$$(13)\qquad (a_1 + a_2 + \ldots + a_m)^n =$$
$$\sum_{i_1, i_2 \ldots i_m} \frac{\Pi(n)}{\Pi(i_1)\,\Pi(i_2)\ldots\Pi(i_m)}\, a_1^{\,i_1} a_2^{\,i_2} \ldots a_m^{\,i_m}$$
$$(i_1, i_2 \ldots i_m = 0, 1, \ldots n;\; i_1 + i_2 + \ldots + i_m = n)$$

der nur eine $(n-1)$fache Summe darstellt, weil $i_1 + i_2 \ldots + i_m = n$ ist.

Besondere Arten von Summen ergeben sich bei der Entwickelung des Produktes

$$(14)\qquad \prod_i (x + x_i) \qquad (i = 1, 2, \ldots n)$$

nach Potenzen von x in der Form einer ganzen Funktion:

$$(15)\quad x^n + f_1(x_1, \ldots x_n)\, x^{n-1} + f_2(x_1, \ldots x_n)\, x^{n-2} + \ldots + f_n(x_1, \ldots x_n),$$

deren Koeffizienten von $x_1, x_2 \dots x_n$ abhängig sind und zwar in der Weise, dafs sie ihre Form nicht ändern, wenn man mit $x_1, \dots x_n$ beliebige Vertauschungen vornimmt. Solche Funktionen nennt man symmetrische, und zwar nennt man im besonderen die hier in der Entwickelung des Produktes auftretenden Funktionen $f_i(x_1 \dots x_n)$ die n symmetrischen Grundformen der Gröfsen $x_1 \dots x_n$ aus einem Grunde, den wir später kennen lernen werden (§ 100).

Es ist leicht festzustellen, dafs

$$(16) \quad f_1(x_1, \dots x_n) = \sum_i x_i, \quad f_n(x_1 \dots x_n) = \prod_i x_i \; (i = 1, 2 \dots n)$$

ist. Die Bildungsweise der allgemeinen Form $f_i(x_1, \dots x_n)$ läfst sich in folgender Weise charakterisieren: Sie ist eine Summe, die aus allen möglichen von einander verschiedenen Produkten von je i der n Gröfsen $x_1, \dots x_n$ gebildet ist, und kann daher in der Form geschrieben werden:

$$(17) \quad f_i(x_1, x_2, \dots x_n) = \sum_{(k_1, k_2 \dots k_i)} x_{k_1} \, x_{k_2} \dots x_{k_i},$$

wo $(k_1, k_2, \dots k_i)$ alle möglichen Kombinationen von je i der Zahlen $1, 2 \dots n$ darstellt, deren Anzahl gleich

$$n_i = \frac{n(n-1) \dots (n-i+1)}{1 \cdot 2 \quad \dots \quad i}$$

ist.

Läfst man alle Gröfsen $x_1, x_2 \dots x_n$ einander gleich werden, so ergiebt sich der Binomialsatz.

---

§ 6. Taylorsche Entwickelung. Ableitungen.

Ersetzt man das Argument x der ganzen Funktion

$$F(x) = \sum_i a_i x^{n-i} \qquad (i = 0, 1 \dots n)$$

durch die Summe $(x + h)$, so geht sie in eine Funktion der beiden Gröfsen x und h über, die wir nach Potenzen von h ordnen können, wo dann die Koeffizienten ganze Funktionen von x sind, und die wir also darstellen können in der Form:

$$F(x+h) = \sum_i \frac{\overset{(i)}{F}(x)}{\Pi(i)} h^i \qquad (i = 0, 1 \dots n).$$

Aus der folgenden Umformung

$$F(x+h) = \sum_{i=0}^{i=n} a_i (x+h)^{n-i} = \sum_{i=0}^{i=n} a_i \sum_{k=0}^{k=n-i} (n-i)_k \, x^{n-i-k} h^k$$
$$= \sum_{k=0}^{k=n} \frac{h^k}{\Pi(k)} \sum_{i=0}^{i=n-k} \frac{\Pi(n-i)}{\Pi(n-i-k)} a_i \, x^{n-i-k}$$

erhält man direkt

$$\overset{(o)}{F}(x) = \sum_{i=0}^{i=n} a_i \, x^{n-i} = F(x)$$

$$\overset{(1)}{F}(x) = \sum_{i=0}^{i=n-1} (n-i) \, a_i \, x^{n-i-1} = F'(x)$$

allgemein

$$\overset{(k)}{F}(x) = \sum_{i=0}^{i=n-k} (n-i)(n-i-1)\ldots(n-i-k+1) \, a_i \, x^{n-i-k}$$
$$(k = 1, \ldots n).$$

Das von h freie Glied in der Entwickelung von $F(x+h)$ stimmt also mit $F(x)$ überein, wie auch schon daraus hervorgeht, dafs $F(x+h)$ für $h=o$ den Wert $F(x)$ annehmen mufs. Die übrigen Funktionen $\overset{(1)}{F}(x)$, $\overset{(2)}{F}(x)$, ... $\overset{(n)}{F}(x)$ nennt man die erste, zweite ... nte Ableitung oder Derivierte von $F(x)$. In der Differentialrechnung werden sie in verallgemeinerter Form als Differentialquotienten*) in Betracht gezogen und bezeichnet durch

$$D_x^i \, F(x) = \frac{d^i F(x)}{dx^i} = \overset{(i)}{F}(x).$$

Sie lassen sich aus einander durch einen einfachen Prozefs (das Differenzieren) bilden. Um die hierbei obwaltenden Bildungsgesetze zu erkennen, beweisen wir zunächst die folgenden Gleichungen

$$\text{(I)} \quad D_x^k \sum_{(i)} F_i(x) = \sum_{(i)} D^k F_i(x)$$

$$\text{(II)} \quad D_x^k (AF(x)) = AD_x^k F(x)$$

in deren letzterer A eine von x unabhängige Gröfse bezeichnen soll. Die erstere ergiebt sich, wenn

$$F(x) = \sum_{(i)} F_i(x)$$

---

*) Vgl. Bd. X. dieser Sammlung.

gesetzt wird, aus der Entwickelung

$$F(x+h) = \sum_{(i)} F_i(x+h) = \sum_{(i,k)} \overset{(k)}{F_i(x)} \frac{h^k}{\Pi(k)}$$

unmittelbar, wenn man den Koeffizienten von $\frac{h^k}{k!}$ in Betracht zieht, während bei der Ableitung der zweiten zu beachten ist, dafs A sich nicht ändert, wenn man x durch $x+h$ ersetzt, so dafs die einfache Multiplikation der Entwickelung von $F(x+h)$ mit A sofort zum Beweise hinführt.

Auf Grund des binomischen Satzes

$$(x+h)^n = \sum_k n(n-1)\ldots(n-k+1)\, x^{n-k} \frac{h^k}{\Pi(k)} \qquad (k=0,1\ldots n)$$

ergeben sich nun sofort die Derivierten von x^n als

$$n x^{n-1},\ n(n-1)\, x^{n-2},\ \ldots$$
$$n(n-1)\ldots 3x^2,\ n(n-1)\ldots 2x,\ n!$$

und wenn man nun (II) beachtet, so geht hervor, dafs hierbei

$$D_x^{i+k}(x^n) = D_x^i D_x^k (x^n)$$

ist. Nunmehr läfst sich zeigen, dafs allgemein

$$\text{(III)} \qquad D_x^{i+k} F(x) = D_x^i D_x^k F(x)$$

ist. Wird nämlich F(x) dargestellt in der Form

$$F(x) = \sum_h a_h x^{n-h} \qquad (h=0,1\ldots n)$$

so ist

$$\begin{aligned} D_x^{i+k} F(x) &= \sum_h D_x^{i+k}(a_h x^{n-h}) = \sum_h a_h D_x^{i+k}(x^{n-h}) \\ &= \sum_h a_h D_x^i D_x^k (x^{n-h}) = \sum_h D_x^i \left(a_h D_x^k (x^{n-h})\right) \\ &= D_x^i \sum_h a_h D^k (x^{n-h}) = D_x^i \sum_h D_x^k (a_h x^{n-h}) \\ &= D_x^i D_x^k \left(\sum_h a_h x^{n-h}\right) = D_x^i D_x^k F(x). \end{aligned}$$

Aus diesem Satze ergiebt sich, dafs

$$\overset{(2)}{F(x)} = D_x \overset{(1)}{F(x)},\quad \overset{(3)}{F(x)} = D_x \overset{(2)}{F(x)}, \ldots \overset{(n)}{F(x)} = D_x \overset{(n-1)}{F(x)}$$

ist, so dafs die Bildung der höheren Ableitungen auf die Bildung der ersten Ableitung einer beliebigen Funktion zurückgeführt ist. Für diese gilt nun noch das Gesetz:

$$\text{(IV)}\quad D_x(A(x)B(x)) = A(x)B'(x) + A'(x)B(x),$$

das man am einfachsten durch die Betrachtung der Entwickelung von $A(x+h)B(x+h)$ beweist, in der die rechte Seite den Koeffizienten von h darstellt. Ist

$$F(x) = (x-x_1)(x-x_2)\ldots(x-x_n),$$

so ergiebt sich durch wiederholte Anwendung dieses Satzes, daſs

$$F'(x) = \sum_{i=1}^{i=n}(x-x_1)\ldots(x-x_{i-1})(x-x_{i+1})\ldots(x-x_n)$$
$$= \sum_{i=1}^{i=n}\frac{F(x)}{x-x_i}$$

und daher

$$F'(x_i) = (x_i-x_1)\ldots(x_i-x_{i-1})(x_i-x_{i+1})\ldots(x_i-x_n)$$

ist. Für alles Weitere möge auf Bd. X dieser Sammlung verwiesen werden.

---

§ 7. Taylorsche Entwickelung von Funktionen mehrerer Variabeln. Partielle Ableitungen.

Die Darlegungen im vorigen Paragraphen lassen sich leicht auf ganze Funktionen von mehreren Variabeln ausdehnen und zwar in folgender Weise:

Ersetzt man die Variabeln $x_1, x_2, \ldots x_n$ einer ganzen Funktion $F(x_1, x_2, \ldots x_n)$ beziehungsweise durch x_1+h_1, x_2+h_2, $\ldots x_n+h_n$, so geht sie in eine Funktion der $2n$ Variabeln $x_1, x_2, \ldots x_n, h_1, h_2, \ldots h_n$ über, die man nach Potenzen von $h_1, h_2, \ldots h_n$ entwickeln kann. Den Koeffizienten von $\frac{h_1^{i_1}\; h_2^{i_2}\;\ldots\; h_n^{i_n}}{\Pi(i_1)\,\Pi(i_2)\ldots\Pi(i_n)}$, der sich als ganze Funktion von $x_1, x_2 \ldots x_n$ darstellt, bezeichnen wir durch

$$D^{(i_1, i_2, \ldots, i_n)}_{(x_1, x_2, \ldots x_n)} F(x_1, x_2, \ldots x_n).$$

Bezeichnet A eine von $x_1, x_2, \ldots x_n$ unabhängige Gröſse, so ergiebt sich:

(I) $D^{(i_1, i_2 \ldots i_n)}_{(x_1, x_2 \ldots x_n)} (AF(x_1 \ldots x_n)) = AD^{(i_1, i_2 \ldots i_n)}_{(x_1, x_2 \ldots x_n)} F(x_1 \ldots x_n).$

Ferner leitet man ab:

(II) $D^{(i_1, i_2 \ldots i_n)}_{(x_1, x_2 \ldots x_n)} (\sum_k F_k(x_1 \ldots x_n)) = \sum_k D^{(i_1, i_2 \ldots i_n)}_{(x_1, x_2 \ldots x_n)} F_k(x_1 \ldots x_n).$

Aus der Entwickelung

$$(x_1 + h_1)^{m_1} (x_2 + h_2)^{m_2} \ldots (x_n + h_n)^{m_n} = \prod_{k=1}^{k=n} (x_k + h_k)^{m_k}$$

$$= \prod_{k=1}^{k=n} \left\{ \sum_{i_k=0}^{i_k=m_k} m_k (m_k - 1) \ldots (m_k - i_k + 1) x_k^{m_k - i_k} \frac{h_k^{i_k}}{\Pi(i_k)} \right\}$$

$$= \sum_{i_1, i_2, \ldots i_n} \prod_{k=1}^{k=n} m_k (m_k - 1) \ldots (m_k - i_k + 1) x_k^{m_k - i_k} \frac{h_k^{i_k}}{\Pi(i_k)}$$

$$(i_1 = 0, 1, \ldots m_1;\ i_2 = 0, 1, \ldots m_2;\ \ldots i_n = 0, 1, \ldots m_n)$$

folgt ferner, dafs

$$D^{(i_1, i_2 \ldots i_n)}_{(x_1, x_2 \ldots x_n)} (x_1^{m_1} x_2^{m_2} \ldots x_n^{m_n}) = D^{i_1}_{x_1} (x_1^{m_1}) D^{i_2}_{x_2} (x_2^{m_2}) \ldots D^{i_n}_{x_n} (x_n^{m_n})$$

ist. Nun ist aber zu beachten, dafs man bei der Bildung der Derivierten nach einer der Variabeln die übrigen als Konstante anzusehen hat. Nach den Sätzen des vorigen Paragraphen kann man daher die rechte Seite auch durch $D^{i_1}_{x_1} D^{i_2}_{x_2} \ldots D^{i_n}_{x_n} (x_1^{m_1} x_2^{m_2} \ldots x_n^{m_n})$ ersetzen. Es ist also

$$D^{(i_1, i_2 \ldots i_n)}_{(x_1, x_2 \ldots x_n)} (x_1^{m_1} x_2^{m_2} \ldots x_n^{m_n}) = D^{i_1}_{x_1} D^{i_2}_{x_2} \ldots D^{i_n}_{x_n} (x_1^{m_1} x_2^{m_2} \ldots x_n^{m_n}).$$

Ähnlich wie es im vorigen Paragraphen geschah, kann man nun leicht unter Anwendung der Gleichungen (I) und (II) zeigen, dafs allgemein

(III) $D^{(i_1, i_2 \ldots i_n)}_{(x_1, x_2 \ldots x_n)} F(x_1, \ldots x_n) = D^{i_1}_{x_1} \ldots D^{i_n}_{x_n} F(x_1, \ldots x_n)$

ist. Die Bildung der Derivierten von Funktionen von mehreren Variabeln ist somit zurückgeführt auf die Bildung der ersten Ableitung nach einer einzigen Variabeln, wobei die übrigen als Konstante betrachtet werden müssen.

---

§ 8. Eigenschaften linearer Formen.

Die linearen Formen sind vor allen übrigen Formen durch einfache Eigenschaften ausgezeichnet, von denen wir jetzt die charakteristischen betrachten wollen:

I) Man kann eine lineare Form multiplizieren, indem man alle Variabeln multipliziert.

Ist nämlich

$$f(x_1, x_2, .. x_n) = a_1 x_1 + a_2 x_2 + .. + a_n x_n,$$

so ist

$$m f(x_1, x_2 \dots x_n) = a_1 (m x_1) + a_2 (m x_2) + \dots + a_n (m x_n)$$
$$= f(m x_1, m x_2, \dots m x_n).$$

Da bei den linearen Formen die Koeffizienten und Variabeln ihre Rollen vertauschen können; so kann man bei der Multiplikation auch die Koeffizienten multiplizieren, die Variabeln dagegen ungeändert lassen.

Übrigens ist dieser Satz nur ein spezieller Fall der oben (§ 3) erwähnten Eigenschaft der Formen, da eine Linearform eine Form ersten Grades ist.

II) Setzt man für die Variabeln $x_1, x_2 .. x_n$ einer Linearform eine Reihe Wertsysteme

$$\begin{matrix} x_1^{(1)} & x_2^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & & x_n^{(2)} \\ \vdots & & & \\ x_1^{(m)} & x_2^{(m)} & \dots & x_n^{(m)}, \end{matrix}$$

so ergiebt sich aus den Gleichungen

$$f(x_1^{(i)}, x_2^{(i)}, \dots x_n^{(i)}) = \sum_k a_k x_k^{(i)} \qquad (i = 1 \dots m;\ k = 1 \dots n)$$

durch Addition:

$$\sum_i f(x_1^{(i)}, x_2^{(i)}, .. x_n^{(i)}) = \sum_{i,k} a_k x_k^{(i)} = \sum_k a_k \sum_i x_k^{(i)}$$
$$= f(\sum_i x_1^{(i)}, \sum_i x_2^{(i)}, \dots \sum_i x_n^{(i)})$$
$$(i = 1 \dots m;\ k = 1 \dots n)$$

woraus man den Satz ablesen kann:

Ersetzt man die n Unbestimmten einer linearen Form durch Summen von gleichviel Gliedern, so läfst sich die lineare Form als Summe von denjenigen Linearformen darstellen, die man erhält, wenn man an Stelle der Summen die Glieder setzt.

---

§ 9. Systeme von linearen Funktionen. Matrizen.

Sind eine Anzahl m von linearen Funktionen, nämlich $f_1, f_2 .. f_m$ gegeben, so ist es von Wichtigkeit, für die Koeffizienten eine übersichtliche Bezeichnungsweise zu haben, die unmittelbar erkennen läfst, in welcher linearen Funktion sie vorkommen und mit welcher Unbestimmten sie dort verbunden sind. Das kann dadurch geschehen, dafs man jeden Koeffizienten mit zwei Indices versieht, von denen der erste übereinstimmt mit dem Index der linearen Funktion, der zweite mit dem der Unbestimmten. Hiernach wollen wir für die m Funktionen folgende Darstellung zu Grunde legen:

$$\begin{aligned} f_1 &= a_{10} + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ f_2 &= a_{20} + a_{21}x_2 + a_{22}x_2 + \ldots + a_{2n}x_n \\ &\vdots \\ f_m &= a_{m0} + a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n. \end{aligned}$$

Ist also i eine der Zahlen 1, 2, .. m, so ist

$$f_i = a_{i0} + a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n$$

und kann kürzer durch

$$f_i = a_{i0} + \sum_k a_{ik}x_k \qquad (i = 1, 2, \ldots m;\ k = 1, 2, \ldots n)$$

oder auch durch

$$f_i = \sum_k a_{ik}x_k \quad (i = 1, 2, \ldots m;\ k = 0, 1, \ldots n;\ x_0 = 1)$$

zur Darstellung kommen, wo zuletzt für x_0 überall der Wert 1 zu setzen ist.

Sieht man bei dem System linearer Funktionen von den Unbestimmten ab, richtet also sein Augenmerk lediglich auf die Koeffizienten und schreibt diese geordnet auf, so erhält man ein rechteckiges System von $(n+1)\,m$ Gröfsen

$$\begin{Vmatrix} a_{10} & a_{11} & a_{12} \dots a_{1n} \\ a_{20} & a_{21} & a_{22} \quad a_{2n} \\ \vdots \\ a_{m0} & a_{m1} & a_{m2} \quad a_{mn} \end{Vmatrix}$$

das man häufig als die vollständige Matrix des Systems von linearen Funktionen bezeichnet. Es hat soviel horizontale Reihen oder Zeilen als Funktionen, aber eine Vertikalreihe oder Spalte (Kolonne) mehr, als Unbestimmte vorhanden sind. Sind die Funktionen homogen, so sind in der ersten Spalte lauter Nullen vorhanden. Läfst man die erste Spalte weg, so erhält man das System

$$\begin{Vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots \\ a_{m1} & a_{m2} \quad a_{mn} \end{Vmatrix},$$

das wir als die verkürzte Matrix bezeichnen wollen, und das bei homogenen Funktionen allein eine Rolle spielt.

---

§ 10. Zusammensetzung linearer Transformationen und Matrizen.

Hat man eine Reihe von m linearen Formen y_1, $y_2, \dots y_m$ der n Unbestimmten $x_1, x_2, \dots x_n$ nämlich

$$y_i = \sum_k a_{ik} x_k, \qquad (i = 1, 2, \dots m;\ k = 1, 2, \dots n)$$

und setzt man an Stelle von $x_1, x_2, \dots x_n$ wieder lineare Formen von r Variabeln $z_1, z_2, \dots z_r$ ein, nämlich

$$x_i = \sum_k b_{ik} z_k, \qquad (i = 1, 2, \dots n;\ k = 1, 2, \dots r)$$

so erhält man das System

$$y_i = \sum_{h,k} a_{ih} b_{hk} z_k = \sum_k \left(\sum_h a_{ih} b_{hk}\right) z_k,$$
$$(i = 1, 2, \dots m;\ k = 1, 2, \dots r;\ h = 1, 2, \dots n)$$

also wieder ein System von m linearen Formen von r Variabeln, dem man die Gestalt

$$y_i = \sum_k c_{ik} z_k \qquad (i = 1, 2, \ldots m;\; k = 1, 2, \ldots r)$$

geben kann, wobei

$$c_{ik} = \sum_h a_{ih} b_{hk} \qquad (i = 1, 2, \ldots m;\; k = 1, 2, \ldots r;\; h = 1, 2, \ldots n)$$

ist. Man kann das Ergebnis kurz so ausdrücken: Zwei lineare Transformationen oder Substitutionen nach einander ausgeführt, lassen sich stets durch eine einzige lineare Transformation ersetzen. Die letztere nennt man auch die aus den beiden anderen zusammengesetzte, wobei wohl auf die Reihenfolge Rücksicht zu nehmen ist, und überträgt diese Bezeichnung, da die Art der Zusammensetzung wesentlich von den Koeffizienten abhängt, und die Variabeln völlig willkürlich sind, auch auf die Matrizen, die wir abgekürzt durch

$$A = (a_{ik}) \qquad (i = 1, 2, \ldots m;\; k = 1, 2, \ldots n),$$
$$B = (b_{ik}) \qquad (i = 1, 2, \ldots n;\; k = 1, 2, \ldots r),$$
$$C = (c_{ik}) \qquad (i = 1, 2, \ldots m;\; k = 1, 2, \ldots r)$$

bezeichnen können. Die beiden ersten, die Komponenten, von den die erste soviel Spalten hat wie die zweite Zeilen, ergeben durch Zusammensetzung (Komposition) das dritte komponierte System, das ebenso viel Zeilen wie die erste, ebenso viel Spalten hat wie die zweite Komponente. Bei der Bildung des komponierten Systems werden die Elemente der einzelnen Zeilen der ersten Matrix (a_{ik}) mit den Elementen der einzelnen Spalten der zweiten Matrix (b_{ik}) multipliziert, und die entstandenen Produkte ergeben dann addiert ein Element des Systems (c_{ik}). Um anzudeuten, dafs durch Zusammensetzung zweier Matrizen A und B die Matrix C entsteht, schreibt man kurz $AB = C$.

Eine Vertauschung der beiden Komponenten ist im allgemeinen nicht gestattet, weil die entstehenden Systeme AB und BA verschieden ausfallen, wovon sich der Leser leicht durch einfache Beispiele überzeugen kann. Dagegen gilt für die Komposition von drei und mehr Matrizen das associative Gesetz, das man kurz in der Form

$$(AB)C = A(BC)$$

ausdrücken kann. Um es zu beweisen, wollen wir annehmen, dafs die Matrizen A, B, C die folgenden sind

$$A = (a_{ik}) \qquad (i = 1, \ldots m; \; k = 1, \ldots n),$$
$$B = (b_{ik}) \qquad (i = 1, \ldots n; \; k = 1, \ldots r),$$
$$C = (c_{ik}) \qquad (i = 1, \ldots r; \; k = 1, \ldots s).$$

Will man die nun Zusammensetzung der drei Matrizen nach der Formel $(AB)C$ ausführen, so ergiebt sich, wenn $AB = D$ und

$$D = (d_{ik}) \qquad (i = 1, 2, \ldots m; \; k = 1, 2, \ldots r)$$

gesetzt wird, zunächst

$$d_{ik} = \sum_h a_{ih} b_{hk}, \qquad (h = 1, \ldots n)$$

die darauf erfolgende Komposition $DC = F$ ergiebt

$$F = (f_{ik}), \qquad (i = 1, \ldots m; \; k = 1, \ldots s)$$

wo

$$f_{ik} = \sum_g d_{ig} c_{gk} = \sum_{h,g} a_{ih} b_{hg} c_{gk}.$$
$$(h = 1, \ldots n; \; g = 1, \ldots r)$$

Dagegen ist bei der Zusammensetzung $A(BC)$ zunächst das System

$$BC = E = (e_{ik}) \qquad (i = 1, \ldots n; \; k = 1, \ldots s)$$

mit

$$e_{ik} = \sum_h b_{ih} c_{hk} \qquad (h = 1, \ldots r)$$

zu bilden, worauf durch Zusammensetzung

$$AE = G = (g_{ik}) \qquad (i = 1, \ldots m; \; k = 1, \ldots s)$$

folgt, wobei

$$g_{ik} = \sum_h a_{ih} e_{hk} = \sum_{h,g} a_{ih} b_{hg} c_{gk},$$
$$(h = 1, \ldots n; \; g = 1, \ldots r)$$

also gleich f_{ik} ist. Daher sind die beiden Matrizen F und G dieselben, und das associative Gesetz ist bewiesen.

---

§ 11. Lineare Transformation quadratischer Formen.

Die soeben gefundene Eigenschaft der linearen Formen, durch lineare Transformation der Variabeln wieder in lineare Formen der neuen Variabeln überzugehen, läfst sich auf Formen höheren Grades übertragen. Es läfst sich leicht zeigen, dafs der Grad einer beliebigen Form durch lineare Transformation der Variabeln nicht geändert wird.

Um dem Leser Gelegenheit zu geben, den Begriff der Komposition zweier Matrizen in einer weiteren Anwendung kennen zu lernen, wollen wir die lineare Transformation der quadratischen Formen betrachten. Eine quadratische Form von n unbestimmten Variabeln $x_1, x_2, \ldots x_n$ läfst sich in der Form darstellen

$$f = \sum_i a_i x_i^2 + \sum_{i,k} b_{ik} x_i x_k,$$

wobei in beiden Summen der Index i die Werte $1, \ldots n$ zu durchlaufen hat, in der zweiten jedoch k nur auf diejenigen Werte zu erstrecken ist, die gröfser als i sind. Man kann jedoch den Ausdruck durch eine Doppelsumme darstellen, in der i und k sämtlich unabhängig von einander die Werte $1, \ldots n$ anzunehmen haben; wenn man nämlich $b_{ik} = 2a_{ik} = 2a_{ki}$ und der Gleichförmigkeit der Bezeichnung wegen $a_i = a_{ii}$ setzt, dann ist

$$f = \sum_{i,k} a_{ik} x_i x_k \qquad (i, k = 1, 2, \ldots n)$$

der Ausdruck für die quadratische Form. Wenn $i \neq k$ ist, so kommt das Produkt $x_i x_k$ zweimal in der Summe vor, nämlich in den beiden Gliedern $a_{ik} x_i x_k$ und $a_{ki} x_i x_k$, so dafs der Koeffizient von $x_i x_k$ gleich $a_{ik} + a_{ki} = 2a_{ik} = b_{ik}$ ist. Ist die quadratische Form in dieser Weise dargestellt, so nennt man die Matrix

$$\begin{pmatrix} a_{11} & a_{12} \ldots a_{1n} \\ a_{21} & a_{22} \ldots a_{2n} \\ \vdots & \\ a_{n1} & a_{n2} \quad a_{nn} \end{pmatrix}$$

die Matrix der quadratischen Form f.

Wenn nun die vorgelegte quadratische Form durch die lineare Transformation

$$x_i = \sum_i \alpha_{ik} x_k' \qquad (i, k = 1, 2, \ldots n)$$

in die transformierte quadratische Form

$$f' = \sum_{i,k} a_{ik}' x_i' x_k' \qquad (i, k = 1, 2, \ldots n)$$

übergeführt wird, so ist

$$a_{ik}' = \sum_{g,h} a_{gh} \alpha_{gi} \alpha_{hk}. \qquad (g, h, i, k = 1, 2, \ldots n)$$

Die Matrix $A' = (a'_{ik})$ kann nun aus der Matrix $A = (a_{ik})$ und der Matrix T der linearen Transformation durch Komposition erzeugt werden, wenn wir noch die aus der letzteren durch gegenseitige Vertauschung der Zeilen und Spalten hervorgehende Matrix $\overline{T}$ einführen, die wir die transponierte Matrix der linearen Transformation nennen wollen. Der Deutlichkeit wegen setzen wir die beiden Matrizen T und $\overline{T}$ ausführlich geschrieben hier her: es ist

$$T = \begin{pmatrix} \alpha_{11} & \alpha_{12} \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{2n} \\ \vdots & & \\ \alpha_{n1} & \alpha_{n2} & \alpha_{nn} \end{pmatrix}, \overline{T} = \begin{pmatrix} \alpha_{11} & \alpha_{21} \ldots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \alpha_{n2} \\ \vdots & & \\ \alpha_{1n} & \alpha_{2n} & \alpha_{nn} \end{pmatrix}.$$

Wie dann eine einfache Rechnung lehrt, ist

$$A' = \overline{T} A T.$$

Wir können das Ergebnis so formulieren:

Man erhält die Matrix der durch eine lineare Transformation hervorgehenden quadratischen Form aus der der ursprünglichen quadratischen Form, indem man sie links mit der transponierten, rechts mit der eigentlichen Matrix der linearen Transformation komponiert.

---

§ 12. Gebrochene lineare Transformationen.

Die Komposition von Matrizen spielt auch eine Rolle bei der Transformation durch gebrochene lineare Funktionen, wenn bei diesen gleiche Nenner vorhanden sind. Setzt man in

$$y_i = \frac{a_{i0} + \sum_k a_{ik} x_k}{a_{00} + \sum_k a_{0k} x_k} \qquad (i = 1 \dots n;\; k = 1 \dots m)$$

an Stelle von x_i

$$x_i = \frac{b_{i0} + \sum_k b_{ik} z_k}{b_{00} + \sum_k b_{0k} z_k}, \qquad (i = 1 \dots m;\; k = 1 \dots r)$$

so ergiebt sich ein System von der Form

$$y_i = \frac{c_{i0} + \sum_k c_{ik} z_k}{c_{00} + \sum_k c_{0k} z_k}, \qquad (i = 1 \dots n;\; k = 1 \dots r)$$

und es ist hierbei

$$(a_{ik})\,(b_{kh}) = (c_{ih}),$$
$$(i = 0, 1, \dots n;\; k = 0, 1, \dots m;\; h = 0, 1 \dots r)$$

wie man durch einfache Ausrechnung, leichter aber daraus erkennt, dafs sich jedes System gebrochener linearer Transformation der genannten Art mit Einführung eines Proportionalitätsfaktors ϱ als ein System ganzer linearer Transformationen auffassen läfst. So kann man das erste System durch das folgende ersetzen

$$\varrho y_i = \sum_k a_{ik}. \qquad (i = 0, 1, \dots n;\; k = 0, 1, \dots r).$$

Das einfachste Beispiel für diesen Fall bildet die Zusammensetzung gebrochener linearer Transformationen einer Veränderlichen. Aus

$$y = \frac{a + bx}{c + dx}, \qquad x = \frac{a' + b'x'}{c' + d'x'}$$

folgt

$$y = \frac{a'' + b''x'}{c'' + d''x'},$$

und hierbei ist

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix}.$$

Jede lineare Transformation kann man zerlegen und zwar in zwei Arten ganzer Transformationen von der Form

$$x' = a + x, \quad x' = a x$$

und eine gebrochene von der Form

$$x' = \frac{1}{x}.$$

Setzt man zwei Transformationen

$$x' = a + x, \quad x' = \frac{1}{x}$$

zusammen, so erhält man die neue Transformation

$$x' = a + \frac{1}{x},$$

die mit den Kettenbrüchen in Zusammenhang steht, zu deren Betrachtung wir jetzt übergehen wollen.

---

§ 13. Kettenbrüche.

Unter einem Kettenbruch versteht man eine Entwickelung von folgender Form

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots + \cfrac{1}{a_i + \cfrac{1}{x_{i+1}}}}}.$$

Die Gröfsen $a_1, a_2, \ldots a_n$ nennt man die Teilnenner, x_{i+1} die Schlufszahl. Wie leicht erkennbar, geht ein solcher Kettenbruch hervor durch eine Reihe von linearen gebrochenen Transformationen, nämlich wenn man setzt

$$x = a_0 + \frac{1}{x_1}, \quad x_1 = a_1 + \frac{1}{x_2}, \ldots \quad x_i = a_i + \frac{1}{x_{i+1}} \ldots$$

Die hierbei auftretenden Matrizen sind

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix}, \ldots \quad \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix}, \ldots$$

Sie unterscheiden sich nur durch die ersten Elemente von einander, die — abgesehen von der ersten Matrix — den Teilnennern des Kettenbruches entsprechen. Die durch Komposition entstehenden Matrizen lassen sich auf folgende Weise bestimmen. Komponiert man zunächst eine beliebige Matrix $\begin{pmatrix} P' & P \\ Q' & Q \end{pmatrix}$ mit $\begin{pmatrix} h & 1 \\ 1 & 0 \end{pmatrix}$, so ergiebt sich

$$\begin{pmatrix} P' & P \\ Q' & Q \end{pmatrix} \begin{pmatrix} h & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P'' & P' \\ Q'' & Q' \end{pmatrix}$$

wobei

$$P'' = P'h + P$$
$$Q'' = Q'h + Q$$

gesetzt ist. Die erste Spalte erscheint somit nach der Komposition als letzte. Daraus folgt, dafs wir allgemein

$$\begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_i & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} P_{i+1} & P_i \\ Q_{i+1} & Q_i \end{pmatrix}$$

setzen können, und dafs

$$P_{i+1} = a_i P_i + P_{i-1}$$
$$Q_{i+1} = a_i Q_i + Q_{i-1}$$

ist. Speziell ist also

$$\begin{matrix} P_0 = 1, & P_0 = a_0, & P_2 = a_0 a_1 + 1 \\ Q_0 = 0, & Q_1 = 1, & Q_2 = a_1 \end{matrix} \qquad P_1 Q_0 - P_0 Q_1 = -1.$$

Aus den beiden vorletzten Gleichungen folgt nun aber, dafs

$$P_{i+1} Q_i - P_i Q_{i+2} = -(P_i Q_{i-1} - P_{i-1} Q_i)$$

ist. Schreibt man diese Gleichungen für die Werte $i = 0, 1, 2 \ldots n - 1$ hin, so ergiebt sich leicht

$$P_n Q_{n-1} - P_{n-1} Q_n = (-1)^n.$$

Wenden wir uns jetzt zu den Transformationen, die der Komposition der Matrizen entsprechen, so erkennt man, dafs allgemein

$$x = \frac{P_{i+1} x_{i+1} + P_i}{Q_{i+1} x_{i+1} + Q_i}$$

ist. Setzt man hierin $x_{i+1} = o$, so ergiebt sich für x der Wert $\frac{P_i}{Q_i}$. Da, wie wir nun sogleich sehen werden, unter

gewissen Voraussetzungen die Schlufszahlen vernachlässigt werden können, um einen angenäherten Wert der Zahl x zu ergeben, so nennt man $\frac{P_i}{Q_i}$ den i-ten **Näherungsbruch**, P_i den i-ten **Näherungszähler**, Q_i den i-ten **Näherungsnenner**. (Wenn die Teilnenner ganze Zahlen, also auch P_i und Q_i solche sind, so sind P_i und Q_i teilerfremd, und der Bruch $\frac{P_i}{Q_i}$ ist in reduzierter Form dargestellt.)

Wenn man eine reelle Zahl x durch einen Kettenbruch darstellen will, so kann man diese Darstellung so einrichten, dafs die Teilnenner $a_1, a_2, \ldots$ ganze positive Zahlen sind. Zu diesem Zwecke wählt man a_0 so, dafs x zwischen a_0 und $a_0 + 1$ liegt, x_1 also eine positive Gröfse wird, bestimmt dann a_1 als die gröfste in x_1 enthaltene positive Zahl und fährt so weiter fort. Da $Q_0 = 0$ ist, so sind die Näherungsnenner von a_0 unabhängig und lauter positive mit dem Index wachsende Zahlen. Da aufserdem Zähler und Nenner teilerfremd sind, so findet das Verfahren einen Abschlufs, wenn x eine rationale Zahl ist. Ist dagegen x irrational, so kann das Verfahren niemals abbrechen. Dann aber müssen die Gröfsen Q_n über jeden Wert hinauswachsen. Aus der Gleichung

$$x = \frac{P_n x_n + P_{n-1}}{Q_n x_n + Q_{n-1}}$$

folgt nun

$$\frac{P_n}{Q_n} - x = \frac{P_n Q_{n-1} - Q_n P_{n-1}}{Q_n (Q_n x_n + Q_{n-1})} = \frac{(-1)^n}{Q_n (Q_n x_n + Q_{n-1})}$$

und hieraus, dafs die Näherungsbrüche

$$\frac{P_1}{Q_1}, \frac{P_3}{Q_3}, \frac{P_5}{Q_5}, \ldots$$

mit ungeradem Index kleiner, die mit geradem Index dagegen

$$\frac{P_0}{Q_0}, \frac{P_2}{Q_2}, \frac{P_4}{Q_4}, \frac{P_6}{Q_6}, \ldots$$

gröfser als x sind. Die rechte Seite der obigen Gleichung ist dem absoluten Betrage nach kleiner als

$$\frac{1}{Q_n Q_{n+1}} < \frac{1}{Q_n^2}.$$

Da aber Q_n über alle Grenzen hinauswächst, so kann $\frac{P_n}{Q_n}$ als rationaler Näherungswert der irrationalen Zahl x angesehen werden.

---

§ 14. Zusammensetzung rationaler Transformationen.

Die im vorhergehenden Paragraphen betrachtete lineare Transformation ist nur ein spezieller Fall der allgemeinen rationalen Transformation, auf die wir nun noch kurz eingehen wollen. Wie wir früher gesehen haben, sind rationale Funktionen von rationalen Funktionen immer wieder rationale Funktionen. Wenn an Stelle des Argumentes x der rationalen Funktion $\varphi(x)$ eine rationale Funktion $\psi(x)$ eingesetzt wird, so wollen wir die entstehende rationale Funktion kurz mit $\varphi\psi(x)$ bezeichnen, oder auch, da es auf das variable Argument nicht ankommt, durch $\varphi\psi$ mit Weglassung der Variabeln. Es möge bemerkt werden, dafs die beiden Funktionen $\varphi\psi$ und $\psi\varphi$ nicht nur begrifflich, sondern auch im allgemeinen der Form nach von einander verschieden sind. Ist $\varphi\psi = \psi\varphi$, so nennen wir die rationalen Transformationen φ, ψ gegen einander vertauschbar.

Wenn mehrere Transformationen nach einander auszuführen sind, so gilt bei der Zusammensetzung das associative Gesetz, d. h. sind φ, ψ, χ, drei rationale Funktionen, so ist $(\varphi\psi)\chi = \varphi(\psi\chi)$. Die Funktion $\varphi\psi\chi$ kann nämlich auf zwei Arten gebildet werden. Nämlich entweder so, dafs zunächst $\varphi\psi(x) = \vartheta(x)$ und dann $\vartheta\chi(x)$, oder dafs zuerst $\psi\chi(x) = \varrho(x)$ und dann $\varphi\varrho(x)$ entsteht. Beide Arten der Auffassung führen nun aber zu demselben Ergebnis. Setzen wir in der Gleichung

$$\varphi\psi(x) = \vartheta(x)$$

an Stelle von x auf beiden Seiten $\chi(x)$ ein, so ergiebt sich auf der linken Seite, weil

$$\psi\chi(x) = \varrho(x)$$

ist, $\varphi\varrho(x)$, auf der rechten einfach $\vartheta\chi(x)$. Damit ist die Richtigkeit mit der Behauptung dargethan, und es ist nun auch nicht schwer, das erhaltene Gesetz auf mehrere rationale Transformationen auszudehnen.

Sind alle zusammenzusetzenden Transformationen dieselben, so kann man in der Abkürzung noch weiter gehen und dieselbe Bezeichnungsweise wie bei den Potenzen anwenden, also $\varphi\varphi(x) = \varphi^2(x)$, $\varphi\varphi^2(x) = \varphi^3(x)$, allgemein

$$\varphi\varphi^{n-1}(x) = \varphi^n(x)$$

setzen. Es läſst sich leicht zeigen, daſs

$$\varphi^a\varphi^b(x) = \varphi^b\varphi^a(x) = \varphi^{a+b}(x)$$
$$(\varphi^a)^b(x) = \varphi^{ab}(x)$$

ist. Für das Weitere möge auf § 35 verwiesen werden, wo genau analoge Verhältnisse gelten und die notwendigen Schlüsse genauer dargelegt sind.

III. Abschnitt.

Teilbarkeit der ganzen Zahlen.

---

§ 15. Begriff der Teilbarkeit.

Durch die Operationen der Addition, Subtraktion und somit auch der Multiplikation (die ja bei ganzen Zahlen als eine Wiederholung von Additionen anzusehen ist) werden aus ganzen Zahlen stets wieder ganze Zahlen erzeugt. Dagegen ist die Division im Gebiete der ganzen Zahlen nur in besonderen Fällen ausführbar. Sind a und b zwei ganze Zahlen und ist $a > b > 0$, so kann man stets zwei ganze Zahlen g und c so bestimmen, dafs

$$a = g b + c, \qquad c < b$$

ist (wobei das Zeichen $<$ die Gleichheit ausschliefst). Subtrahiert man nämlich b so oft von a, als es möglich ist, also bis ein Rest c erscheint, der kleiner ist als b, so ist g die Anzahl der nötigen Subtraktionen. Man drückt sich gewöhnlich so aus, dafs sich bei der „Division von a durch b" g als Quotient, c als Rest ergiebt. Nur wenn dieser Rest gleich Null ist, ergiebt die Division von a durch b als Resultat eine ganze Zahl, nämlich g. Man drückt diesen Sachverhalt dann dadurch aus, dafs man von a sagt, dafs es durch b *teilbar* oder ein *Vielfaches* (g-faches) von b sei, von b, dafs es *Teiler* von a, in ihm als *Faktor enthalten* sei oder in a *aufgehe*.

Aus dem Begriffe der Teilbarkeit ergeben sich leicht folgende *Fundamentalsätze:*

1) *Eine Summe ist durch jeden gemeinschaftlichen Teiler seiner Summanden teilbar.* 2) *Eine*

Differenz ist durch jeden gemeinschaftlichen Teiler von Minuend und Subtrahend teilbar. 3) Überhaupt ist jede algebraische Summe durch jeden gemeinschaftlichen Teiler seiner Glieder teilbar. 4) Ein Produkt ist durch jeden Teiler eines seiner Faktoren teilbar.

Die Beweise dieser Sätze sind sehr einfach. Sind a und b zwei Zahlen mit einem gemeinschaftlichen Teiler m, so giebt es zwei ganze Zahlen a′ und b′, so dafs $a = a'm$, $b = b'm$ ist. Aus diesen Gleichungen folgt aber

$$a + b = a'm + b'm = (a' + b')m$$
$$a - b = a'm - b'm = (a' - b')m,$$

und hierdurch wird, da $a' + b'$ und $a' - b'$ ganze Zahlen sind, ausgedrückt, dafs $a + b$ und $a - b$ durch m teilbar sind, wie es die beiden ersten Sätze behaupten. Da ein Aggregat sich aus einer Reihe von aufeinanderfolgenden Additionen und Subtraktionen zusammensetzt und ein Produkt als eine Summe von lauter gleichen Summanden auffassen läfst, so ergiebt sich aus den beiden ersten Sätzen auch die Richtigkeit der beiden letzten.

---

§ 16. Eigentliche und uneigentliche Teiler. Primzahlen und zusammengesetzte Zahlen.

Wenn die ganze Zahl a durch b teilbar ist, so ist die dann eindeutig bestimmte Zahl $c = a : b$ ebenfalls ein Teiler von a und wird als der zu b komplementäre Teiler bezeichnet. Läfst man die Zahl 0, die offenbar durch jede ganze Zahl teilbar ist, von vornherein aufser Betracht, so ist jeder Teiler einer Zahl höchstens (an absolutem Betrag, wenn man auch negative Zahlen berücksichtigt) dieser Zahl gleich, weil der komplementäre Teiler mindestens den Betrag 1 hat. Sind zwei Zahlen gegenseitig durch einander teilbar, so sind sie dem Betrage nach gleich. Da alle ganzen Zahlen durch die Einheiten $+1$ und -1 und sich selbst teilbar sind, so kann man diese Teiler als uneigentliche etwaigen andern eigentlichen Teilern gegenüberstellen. Dann ist jeder eigentliche Teiler einer Zahl kleiner als diese und

gröfser als die Einheit (dem Betrage nach). Da es nur eine endliche Menge von ganzen Zahlen giebt, die diese letztere Bedingung erfüllen, so ist man imstande, durch eine beschränkte Anzahl von Versuchen alle Teiler einer gegebenen Zahl zu finden, indem man die Reste bestimmt, die sich durch Division mit allen kleineren Zahlen ergeben, wobei die durch den Rest 0 ausgezeichneten als Teiler herausfallen. Man braucht jedoch, wenn man so mit den kleinsten Zahlen beginnt und zu jeder sich als Teiler erweisenden Zahl den komplementären Teiler mit in Betracht zieht, nur soweit fortzufahren, bis sich der komplementäre Teiler, der ja beständig abnimmt, kleiner als der Teiler erweist. Wenn t der kleinste Teiler von a ist, zu dem der komplementäre Teiler $t' < t$ gehört, so braucht man die Versuche nur bis t' auszudehnen. Aus den Gleichungen

$$a = tt', \quad t' < t$$

folgt aber

$$t'^2 < a < t^2.$$

Man kann daher die Grenze, bis zu der man die Teilbarkeitsversuche zu erstrecken braucht, auch dadurch charakterisieren, dafs ihr Quadrat die gröfste Quadratzahl unter a ist, oder dafs sie selbst die gröfste in $\sqrt{a}$ enthaltene ganze Zahl ist.

Jede Zahl, die keine eigentlichen Teiler besitzt, nennt man eine Primzahl; alle übrigen (nach Ausschlufs von Null und den Einheiten) heifsen zusammengesetzte Zahlen. Jede zusammengesetzte Zahl hat also mindestens einen eigentlichen Teiler und läfst sich als Produkt von mindestens zwei Zahlen darstellen, die beide eigentliche Teiler sind. Sobald diese wieder zusammengesetzte Zahlen sind, lassen sie sich wieder zerlegen, und man kann so in der Zerlegung fortfahren, bis man auf Zahlen stöfst, die keine eigentlichen Teiler mehr besitzen, also Primzahlen sind. Da dies aber wegen der abnehmenden Gröfse des Teiles eintreten mufs, so kann man jede zusammengesetzte Zahl als ein Produkt von lauter Primzahlen darstellen, wobei man mehrfach vorkommende Primzahlen zu Primzahlpotenzen vereinigen kann. Dafs diese Darstellung aber, abgesehen von der Reihenfolge der Faktoren, eine völlig bestimmte ist,

geht aus dieser Ableitung, die für die Art der Zerlegung der Willkür Spielraum läfst, nicht hervor. Wir werden den Beweis hierfür später (§ 23) erbringen und dann aber auch darlegen, wie sich das Problem der Bestimmung aller Teiler einer gegebenen Zahl leichter lösen läfst.

Es giebt eine unbeschränkte Menge von Primzahlen, da man zu jeder Primzahl p eine noch gröfsere auf folgende Weise ableiten kann. Bedeutet nämlich m das Produkt aller Primzahlen 2, 3, 5 .... p, die nicht gröfser als p sind, a eine beliebige Zahl, so ist $am + 1$ durch keine der genannten Primzahlen teilbar, weil sich bei der Division mit jeder der Rest 1 ergiebt. Zerlegt man also $am + 1$, wenn es nicht selbst schon eine Primzahl darstellt, in ein Produkt von Primzahlen, so ist jede derselben gröfser als p.

Wegen der aufserordentlichen Rolle, die die Primzahlen in der Zahlentheorie spielen, hat man es für nützlich gehalten, Primzahltafeln herzustellen, die alle Primzahlen bis zu einer gewissen Grenze enthalten. Das einfachste Verfahren, solche Tafeln zu konstruieren, besteht darin, dafs man zunächst alle Zahlen bis zu der gewünschten Grenze aufschreibt, darauf alle durch 2, 3, 5 u. s. w. teilbaren Zahlen entfernt, was selbstverständlich auch schon beim Aufschreiben geschehen kann. Hat man dies Verfahren mit allen Primzahlen 2, 3 .. p durchgeführt, und ist q die auf p folgende, nicht entfernte Zahl, so ist q eine Primzahl, und man hat dann alle durch q teilbaren Zahlen zu entfernen, und so fortzufahren.

Beispiele. Um 857 in Bezug auf die Zerlegbarkeit zu untersuchen, stellt man zunächst fest, dafs $29^2 < 857 < 30^2$ ist und findet bei der Division von 857 durch die Primzahlen 2, 3, 5, 7, 11, 13, 17, 19, 23, 29 die Reste 1, 2, 2, 3, 10, 12, 7, 2, 6, 16; daher ist 857 eine Primzahl und nicht zerlegbar. Dagegen ergeben sich bei 1081, wo $32^2 < 1081 < 33^2$ ist, bei der Division durch 2, 3, 5, 7, 11, 13, 19 die Reste 1, 1, 1, 3, 3, 2, 17, während bei der Division mit der auf 19 folgenden Primzahl 23 sich die Zerlegung $1081 = 23 \,.\, 47$ ergiebt.

§ 17. Gröfster gemeinschaftlicher Teiler mehrerer Zahlen. Modulsysteme.

Da mehrere Zahlen a, b, c, ... stets einen gemeinschaftlichen Teiler, nämlich die Einheit, besitzen und ferner die Anzahl der gemeinschaftlichen Teiler nur eine beschränkte sein kann, so mufs unter ihnen ein gröfster vorhanden sein, und das zunächst sich darbietende und begrifflich einfachste Verfahren, diesen zu ermitteln, würde einfach darin bestehen, zunächst von jeder der Zahlen a, b, c, ... nach der im vorigen Paragraphen angegebenen Methode die sämtlichen Teiler aufzusuchen, darauf alle gemeinschaftlichen Teiler auszusondern, um dann schliefslich zuzusehen, welcher der gröfste ist. Diesen gröfsten gemeinschaftlichen Teiler der Zahlen a, b, c, ... bezeichnen wir durch das Zeichen

$$(a, b, c, \ldots)$$

also dadurch, dafs wir die Zahlen a, b, c, .. durch Kommata getrennt nebeneinanderschreiben und durch eine Klammer umschliefsen. Der Nutzen und die Wichtigkeit einer solchen Bezeichnung wird sich bald von selbst herausstellen. Wir nennen den gröfsten gemeinschaftlichen Teiler von a, b, c, ... auch das aus den Zahlen a, b, c, .. als Elementen gebildete Modulsystem. Ist $(a, b, c, ..) = 1$, so haben a, b, c, .. aufser der Einheit keinen gemeinschaftlichen Teiler. Haben zwei Zahlen a, b aufser der Einheit keinen gemeinschaftlichen Teiler, so nennt man sie teilerfremd, relativ prim, relative Primzahlen. Dafs a durch b teilbar ist, kann man kurz durch $(a, b) = b$ ausdrücken. Ist p eine Primzahl, a eine beliebige Zahl, so kann (a, p) nur die Werte 1 oder p haben, den letzteren nur, wenn a durch p teilbar, also $a \geqq p$ ist, so dafs bei $a < p$ stets $(a, p) = 1$ ist. Sind p und q verschiedene Primzahlen, so ist stets $(p, q) = 1$. Man erkennt auch, dafs zwei verschiedene Primzahlpotenzen stets teilerfremd sind. Dagegen ist $(p^a, p^b) = p^c$, wenn c die kleinere der beiden Zahlen a und b ist. Zwei Modulsysteme, die den gleichen Wert darstellen, nennt man auch äquivalent.

Über die Modulsysteme gilt nun folgender Fundamentalsatz:

Ein Modulsystem bleibt seinem Werte nach ungeändert, wenn man irgend eins seiner Elemente um ein beliebiges andere vermehrt oder vermindert.

Um zu beweisen, dafs

$$(a \pm b, b, c, \ldots) = (a, b, c, \ldots)$$

ist, benutzen wir die beiden Sätze (1) und (2) in § 15. Nach ihnen mufs nämlich jeder gemeinschaftliche Teiler von a und b auch $a \pm b$ teilen, und daher der gröfste gemeinschaftliche Teiler von a, b, c, ... ein gemeinschaftlicher Teiler von $a \pm b, b, c, \ldots$ sein, so dafs $(a, b, c, \ldots) \leqq (a \pm b, b, c, \ldots)$ ist. Andererseits mufs aber auch jeder gemeinschaftliche Teiler von $a \pm b$ und b auch $(a \pm b) \mp b = a$ teilen, also der gröfste gemeinschaftliche Teiler von $(a \pm b, b, c, \ldots)$ ein gemeinschaftlicher Teiler von a, b, c, .., somit $(a \pm b, b, c, \ldots) \leqq (a, b, c, \ldots)$ sein. Aus den beiden Beziehungen ergiebt sich nun sofort die Behauptung.

Wendet man nun diesen Satz wiederholt auf ein Modulsystem

$$(a_0, a_1, a_2, \ldots a_n)$$

an, so ergiebt sich, dafs dieses übergeführt werden kann in

$$(a_0 + g_1 a_1 + g_2 a_2 + \ldots + g_n a_n, a_1, a_2, \ldots a_n),$$

wo $g_1, g_2, \ldots g_n$ ganze positive oder negative Zahlen sind, und so erhalten wir den Satz:

Ein Modulsystem bleibt dem Werte nach ungeändert, wenn man eins seiner Elemente um eine lineare Form der übrigen mit ganzzahligen Koeffizienten vermehrt.

Nehmen wir $a_0 = 0$ an und bedenken, dafs *die Null jedem beliebigen Modulsystem als Element beigefügt werden darf*, weil sie durch alle Zahlen teilbar ist, so dafs also

$$(a_1, a_2, \ldots a_n, 0) = (a_1, a_2, \ldots a_n)$$

ist, so ergiebt sich mit der Gleichung

$$(g_1 a_1 + g_2 a_2 + \ldots + g_n a_n, a_1, a_2, \ldots a_n) = (a_1, a_2, \ldots a_n),$$

dafs man jedem Modulsystem eine lineare Form seiner Elemente mit ganzzahligen Koeffizienten hinzufügen kann.

Hieraus aber ergiebt sich nun der folgende allgemeine Satz:

Wenn die Elemente zweier Modulsysteme in einem solchen Abhängigkeitsverhältnisse zu einander stehen, dafs die Elemente des einen als lineare Formen der Elemente des andern und umgekehrt sich darstellen lassen, so sind sie äquivalent.

Denn sind $(a_1, a_2, .. a_n)$ und $(b_1, b_2, .. b_n)$ solche Modulsysteme, so ist

$$(a_1, a_2, .. a_n) = (a_1, a_2 ... a_n, b_1, b_2, ... b_m)$$
$$(b_1, b_2, .. b_m) = (b_1, b_2 .. b_m, a_1, a_2 .. a_n),$$

da man nach dem letzten Satze dem System $(a_1 .. a_n)$ die Gröfsen $b_1, b_2 .. b_m$ als lineare Formen von $a_1 .. a_n$, dem System $(b_1, b_2 .. b_m)$ die Gröfsen $a_1, a_2 .. a_n$ als lineare Formen von $b_1, b_2 ... b_m$ hinzufügen kann. Aus den beiden Gleichungen folgt dann aber

$$(a_1, a_2, .. a_n) = (b_1, b_2, .. b_m).$$

---

§ 18. Bestimmung des gröfsten gemeinschaftlichen Teilers. Zusammenhang der Elemente äquivalenter Modulsysteme.

Die soeben entwickelten Sätze können dazu benutzt werden, ein Modulsystem zu berechnen oder anders ausgedrückt, den gröfsten gemeinschaftlichen Teiler mehrerer Zahlen zu finden. Das hierbei anzuwendende Reduktionsverfahren findet sich in der speziellen Form für zwei Zahlen schon in den Elementen des Euklid und wird daher häufig als Euklidischer Algorithmus bezeichnet. Wir wollen diesen am häufigsten vorkommenden Fall zunächst betrachten.

Sind a und b die beiden Zahlen, deren gröfster gemeinschaftlicher Teiler bestimmt werden soll, und nimmt man, wie es gestattet ist, an, dafs $a > b > o$ ist, so kann man (§ 15) zwei ganze Zahlen g und c so ermitteln, dafs

$$a = gb + c, \quad c < b$$

ist. Dann aber ergiebt sich aus den Sätzen des vorigen Paragraphen

$$(a, b) = (g b + c, b) = (b, c),$$

so dafs die Bestimmung von (a, b) auf (b, c) zurückgeführt ist, wo kleinere Zahlen als Elemente auftreten. Ist nun $c > 0$, so kann man auf (b, c) dasselbe Reduktionsverfahren anwenden, wobei es durch (c, d) ersetzt wird, wo $d < c$ ist. So kann man weiter fortfahren, so lange noch das kleinere Element von Null verschieden ist. Da aber die Zahlen

$$a, b, c, d, \ldots$$

immer kleiner werden, so mufs nach einer bestimmten Anzahl von Reduktionen dieser Fall eintreten und das kleinere Element gleich 0 werden. Nennt man das vorhergehende Glied der Reihe t, so ist

$$(a, b) = (b, c) = (c, d) = \ldots = (t, 0) = t$$

und somit t der gröfste gemeinschaftliche Teiler von a und b.

Ähnlich kann man bei der Aufsuchung des gröfsten gemeinschaftlichen Teilers mehrerer Zahlen $a_1, a_2, \ldots a_n$ verfahren. Bezeichnen wir durch a eine beliebige von ihnen, die gröfser als eine andere b ist, so führt die Zerlegung

$$a = g b + a', \quad a' < b$$

auf ein neues Modulsystem, in dem a durch die kleinere Zahl a' ersetzt ist. Fährt man so fort, so mufs zuletzt der Fall eintreten, dafs das Modulsystem nur noch eine einzige von 0 verschiedene Zahl t enthält, die dann den gröfsten gemeinschaftlichen Teiler $t = (a_1, a_2, \ldots a_n)$ darstellt.

Richten wir nun unser Augenmerk auf die einzelnen Modulsysteme, die durch die soeben angegebene Reduktionsmethode entstehen, so können wir sagen, dafs die Elemente jedes vorangehenden mit denen des folgenden wie auch umgekehrt durch eine lineare Transformation mit ganzzahligen Koeffizienten — sehr spezieller Art, da immer nur ein Element wirklich linear transformiert wird — zusammenhängen. Da nun aber aus der Zusammensetzung linearer Transformationen stets wieder solche hervorgehen (§ 10), und zwar auch wieder mit ganzzahligen Koeffizienten, so sind überhaupt irgend zwei der bei der Reduktion auftretenden Modulsysteme durch eine solche lineare Trans-

formation wechselseitig mit einander verbunden. Wenden wir dies speziell auf das ursprüngliche System $(a_1\ a_2, \ldots a_n)$ und das aus der einzigen Zahl t bestehende Schlufssystem an, so ergiebt sich, dafs man 2n ganze positive oder negative Zahlen $g_1, g_2, \ldots g_n, h_1, h_2, \ldots h_n$ so bestimmen kann, dafs

$$t = \sum_i g_i a_i, \qquad a_i = h_i t \qquad (i = 1, 2 \ldots n)$$

ist. Nehmen wir ferner an, dafs das Modulsystem $(b_1\ b_2 \ldots b_m)$ dem Werte nach dem Modulsystem $(a_1, a_1, . a_n) = t$ äquivalent ist, so müssen sich auch seine Elemente linear durch t, wie umgekehrt dieses linear und homogen durch seine Elemente ausdrücken lassen. Da man also von $(a_1, a_2, \ldots a_n)$ sowohl wie von $(b_1, b_2, \ldots b_m)$ zu t wie auch umgekehrt durch eine lineare Transformation mit ganzen Koeffizienten übergehen kann, so ergiebt sich zum Schlufssatz des vorigen Paragraphen als Umkehrung der Satz:

Die Elemente zweier äquivalenter Modulsysteme sind durch einander wechselseitig als lineare Formen mit ganzzahligen positiven oder negativen Koeffizienten darstellbar.

Fassen wir die beiden Sätze zusammen, so können wir sagen:

Die hinreichenden und notwendigen Bedingungen dafür, dafs zwei Modulsysteme äquivalent sind, bestehen darin, dafs wechselseitig die Elemente des einen sich als lineare Formen der Elemente des anderen mit ganzen positiven oder negativen Koeffizienten darstellen lassen.

---

§ 19. Multiplikation der Modulsysteme.

Den zuletzt abgeleiteten wichtigen Satz wollen wir nun sogleich zur Anwendung bringen. Nehmen wir an, dafs

$$a = (a_1, a_2, \ldots a_n)$$

ist, so lassen sich 2n ganze positive oder negative Zahlen $g_1, \ldots g_n;\ h_1, \ldots h_n$ so bestimmen, dafs

$$a = \sum_i g_i a_i, \quad a_i = h_i a \qquad (i = 1, 2, \ldots n)$$

ist. Durch Multiplikation dieser Gleichungen mit einer ganzen Zahl b ergiebt sich nun aber

$$ab = \sum_i g_i \,.\, a_i b, \quad a_i b = \sum_i h_i \,.\, ab. \quad (i = 1, 2, \ldots n)$$

Hierdurch wird nun aber ausgedrückt, dafs $(a_1 b, a_2 b, .. a_n b)$ äquivalent ab ist. Da $a = (a_1, a_2 .. a_n)$ war, so ist also

$$(a_1, a_2, \ldots a_n)\, b = (a_1 b, a_2 b, \ldots a_n b),$$

und man erhält demnach den Satz:

Man kann ein Modulsystem mit einer ganzen Zahl multiplizieren, indem man jedes Element mit ihr multipliziert und die erhaltenenen Produkte zu einem Modulsystem vereinigt.

Liest man die Formel, indem man von der rechten Seite ausgeht und die linke als eine Transformation der rechten betrachtet, so kann ihr Inhalt auch so ausgedrückt werden:

Enthalten die Elemente eines Modulsystems einen gemeinsamen Faktor, so kann man es als ein Produkt darstellen, dessen einer Faktor der gemeinschaftliche Faktor der Elemente ist, während der andere Faktor das Modulsystem ist, das aus den übrigen Faktoren gebildet werden kann.

Betrachten wir nun zwei Modulsysteme

$$a = (a_1, a_2, \ldots a_n)$$
$$b = (b_1, b_2, \ldots b_m)$$

so ergiebt sich durch zweimalige Anwendung des ersten Satzes

$$\begin{aligned} ab &= (a_1, a_2, \ldots a_n)\, b = (a_1 b, a_2 b, \ldots a_n b) \\ &= [a_1 (b_1, b_2, \ldots b_m), a_2 (b_1, b_2, .. b_m), \ldots a_n (b_1, b_2, \ldots b_m)] \\ &= (a_1 b_1, \ldots a_1 b_2, a_1 b_m, a_2 b_1, a_2 b_2, \ldots a_2 b_m, \ldots a_n b_1, a_n b_2, \ldots a_n b_m), \end{aligned}$$

sodafs also

$$(a_1, a_2, \ldots a_n)(b_1, b_2, \ldots b_m) = (a_i b_k) \quad (i, k = 1, 2, \ldots n)$$

ist, wo das auf der rechten Seite stehende System durch eine abgekürzte Schreibweise hinreichend deutlich dargestellt ist, und damit folgt der Satz:

Man kann das Produkt zweier Modulsysteme auch dadurch bilden, dafs man jedes Element des einen mit jedem Element des anderen multipliziert und die erhaltenen Produkte zu einem Modulsystem vereinigt.

---

§ 20. Zerlegung einer Zahl in Primfaktoren.

Von den mannigfachen Folgerungen, die sich aus den letzten Sätzen ziehen lassen, wollen wir hier nur diejenigen betrachten, aus denen sich die Bestimmtheit der Zerlegung jeder Zahl in ein Produkt von Primzahlen ableiten, also die in § 16 erwähnte Lücke ausfüllen läfst. Wir schicken hierzu einige Betrachtungen vorauf:

Man beweist zunächst leicht die Richtigkeit der folgenden Äquivalenz

$$(a, b, c)\,(bc, ca, ab) = (b, c)\,(c, a)\,(a, b),$$

indem man auf beiden Seiten den letzten Satz des vorigen Paragraphen zur Anwendung bringt, wodurch sich das Modulsystem

$$(a^2b,\ a^2c,\ b^2a,\ b^2c,\ c^2a,\ c^2b,\ abc)$$

ergiebt. Nimmt man nun $(a, b) = 1$ an, so wird $(a, b, c) = 1$, $(bc, ca, ab) = [c\,(a, b), ab] = (a\,b, c)$, und man erhält, wenn man noch für a, b, c resp. a_1, a_2, b einführt, den folgenden Satz:

Sind a_1 und a_2 teilerfremd, so ist

$$(a_1\,a_2, b) = (a_1, b)\,(a_2, b).$$

Dieser läfst sich nun beträchtlich verallgemeinern: Sind von den Zahlen $a_1, a_2 \ldots a_n$, immer je zwei teilerfremd, so gilt die Äquivalenz

$$(a_1\,a_2 \ldots a_n, b) = (a_1, b)\,(a_2, b) \ldots (a_n, b).$$

Den Beweis führen wir durch vollständige Induktion. Wir nehmen an, dafs sie für $(n-1)$ Zahlen $a_1, a_2, \ldots a_{n-1}$ gelte, dafs also

$$(a_1\,a_2 \ldots a_{n-1}, b) = (a_1, b)\,(a_2, b) \ldots (a_{n-1}, b)$$

ist. Dann ergiebt sich, wenn wir für b die Zahl a_n setzen, zunächst, dafs

$$(a_1\, a_2 \dots a_{n-1}, a_n) = (a_1, a_n)\,(a_2, a_n) \dots (a_{n-1}, a_n) = 1,$$

also $a_1\, a_2 \dots a_{n-1}$ zu a_n teilerfremd ist. Daher ist weiter

$$(a_1\, a_2 \dots a_{n-1}\, a_n, b) = (a_1\, a_2 \dots a_{n-1}, b)\,(a_n, b).$$

Setzen wir nun für den ersten Faktor der rechten Seite das Produkt $(a_1, b) \dots (a_{n-1}, b)$, so erhalten wir in der That die zu beweisende Äquivalenz für n Zahlen $a_1, a_2 \dots a_n$, die also allgemein gültig ist.

Wie wir nun früher gesehen haben, läfst sich jede Zahl m als ein Produkt von lauter Primzahlen darstellen. Angenommen

$$m = p_1^{e_1} p_2^{e_n} \dots p_n^{e_n}$$

sei eine solche Darstellung, bei der $p_1, p_2 \dots p_n$ von einander verschiedene Primzahlen, $e_1, e_2, \dots e_n$ die bezüglichen Grade ihrer Vielfachheit sind, setzen wir nun

$$a_1 = p_1^{e_1},\ a_2 = p_2^{e_2}, \dots a_n = p_n^{e_n},$$

so sind von den Zahlen $a_1, a_2 \dots a_n$ immer je zwei ohne gemeinschaftlichen Teiler und also die Voraussetzungen des soeben bewiesenen Satzes erfüllt. Daher ist

$$(m, b) = (p_1^{e_1}, b)\,(p_2^{e_2}, b) \dots (p_n^{e_n}, b).$$

Wir ersetzen nun b durch irgend eine Primzahlpotenz p^e. Nehmen wir an, dafs p von $p_1, p_2, \dots p_n$ verschieden ist, so ist jeder Faktor auf der rechten Seite gleich 1, und es folgt aus

$$(m, p^e) = 1,$$

dafs m durch keine solche Primzahlpotenz teilbar sein kann. Ist dagegen p gleich einer der Primzahlen $p_1, p_2, \dots p_n$ etwa gleich p_i, so sind alle Faktoren auf der rechten Seite gleich 1 mit Ausnahme des einzigen

$$(p_i^{e_i}, p_i^e),$$

dessen Wert gleich p_i^f ist, wo f die kleinere der beiden Zahlen e, e_i bedeutet, und somit ist dann

$$(m, p_i^e) = p_i^f, \quad f \leqq e, \quad f \leqq e_i.$$

Daher ist m durch p_i^e nur dann teilbar, wenn $e \leqq e_i$ ist, und $p_i^{e_i}$ ist also die höchste Potenz von p_i, durch die

m teilbar ist, und als solche völlig bestimmt. Da also bei einer anders ausgeführten Zerlegung von m keine anderen Primzahlen als $p_1, p_2, \ldots p_n$ und auch nicht mit höheren Exponenten als bezw. $e_1, e_2 \ldots e_n$ auftreten können, so ist die Zerlegung der Zahlen in ein Produkt von Primzahlpotenzen nur auf eine einzige Weise möglich, abgesehen natürlich von der Reihenfolge, in der man sie anordnen kann. Würde bei einer zweiten Darstellung eine der Primzahlen p_i in einer geringeren Potenz als $p_i^{e_i}$ erscheinen, so würde der Annahme entgegen m nicht durch $p_i^{e_i}$ teilbar sein können.

---

§ 21. Anwendungen der Zerlegung in Primfaktoren.

Setzt man die Zerlegung der Zahlen in Primfaktoren voraus, so lassen sich eine Reihe von Aufgaben, von denen wir die meisten schon vorher behandelt haben, in sehr einfacher Weise lösen.

1) Die hinreichende und notwendige Bedingung dafür, daſs eine Zahl a durch eine andere b teilbar ist, besteht darin, daſs sie durch jede in b enthaltene Primzahlpotenz teilbar ist.

Daſs die Bedingung eine notwendige ist, ist keines Beweises bedürftig. Um zu beweisen, daſs sie auch hinreicht, nehmen wir an, daſs

$$b = p_1^{e_1} p_2^{e_2} \ldots p_n^{e_n}$$

die Zerlegung von b, und

$$(a, p_1^{e_1}) = p_1^{e_1}, \ (a, p_2^{e_2}) = p_2^{e_2} \ldots (a, p_n^{e_n}) = p_n^{e_n}$$

sei, dann ergiebt sich aus

$$(a, b) = (a, p_1^{e_1} p_2^{e_2} \ldots p_n^{e_n}) = (a, p_1^{e_1}) (a, p_2^{e_2}) \ldots (a, p_n^{e_n})$$
$$= p_1^{e_1} p_2^{e_2} \ldots p_n^{e_n} = b$$

die Behauptung. Wenn also a durch b teilbar sein soll, so muſs jede in b enthaltene Primzahl bei der Zerlegung von a in mindestens ebenso hoher Potenz wie bei der von b vorkommen.

2) Alle Teiler einer Zahl

$$m = p_1^{e_1} p_2^{e_2} \ldots p_n^{e_n}$$

sind enthalten in der Formel

$$t_{i_1 i_2 \ldots i_n} = p_1^{i_1} p_2^{i_2} \ldots p_n^{i_n},$$

wo i_1 die Werte $0, 1, \ldots e_1$, i_2 die Werte $0, 1, \ldots e_2$, $\ldots i_n$ die Werte $0, 1 \ldots e_n$ annehmen kann. Die Anzahl sämtlicher Teiler ist somit gleich $(e_1 + 1)(e_2 + 1) \ldots (e_n + 1)$. Die Summe der Teiler ergiebt sich durch folgende Rechnung:

$$\sum_{i_1, i_2 \ldots i_n} t_{i_1 i_2 \ldots i_n} = \sum_{i_1, i_2 \ldots i_n} p_1^{i_1} p_2^{i_2} \ldots p_n^{i_n} = \sum_{i_1} p_1^{i_1} \sum_{i_2} p_2^{i_2} \ldots \sum_{i_n} p_n^{i_n}$$

$$= \frac{p_1^{e_1+1} - 1}{p_1 - 1} \cdot \frac{p_2^{e_2+1} - 1}{p_2 - 1} \ldots \frac{p_n^{e_n+1} - 1}{p_n - 1} = \prod_i \frac{p_i^{e_i+1} - 1}{p_i - 1},$$

$$(i = 1, 2, \ldots n)$$

ihr Produkt erhält man aus

$$\prod_{i_1 i_2 \ldots i_n} t_{i_1 i_2 \ldots i_n} = \prod_{i_1 i_2 \ldots i_n} p_1^{i_1} p_2^{i_2} \ldots p_n^{i_n} = \prod_{i_1} p_1^{i_1} \prod_{i_2} p_2^{i_2} \ldots \prod_{i_n} p_n^{i_n}$$

$$= p_1^{\sum_{i_1} i_1} p_2^{\sum_{i_2} i_2} \ldots p_n^{\sum_{i_n} i_n} = p_1^{\frac{e_1(e_1+1)}{2}} p_2^{\frac{e_2(e_2+1)}{2}} \ldots p_n^{\frac{e_n(e_n+1)}{2}}$$

$$= \prod_i p_i^{\frac{e_i(e_i+1)}{2}} \qquad (i = 1, 2, \ldots n).$$

Bei der ersten Rechnung hat man die Summenformel der geometrischen Reihen, bei der zweiten die der arithmetischen Reihen anzuwenden.

3) Um den gröfsten gemeinschaftlichen Teiler von mehreren Zahlen, deren Zerlegung in Primfaktoren ausgeführt ist, zu ermitteln, sucht man die in allen Zerlegungen vorkommenden Primzahlen auf und bestimmt sodann von jeder die niedrigste Potenz, in der sie auftritt. Das Produkt aller so erhaltenen Primzahlpotenzen ist dann der gröfste gemeinschaftliche Teiler der gegebenen Zahlen. Sind keine gemeinschaftlichen Primzahlen vorhanden, so ist die Einheit der einzige gemeinschaftliche Teiler.

Der Beweis für diese Regel ergiebt sich daraus, dafs jeder gemeinschaftliche Teiler der Zahlen nur solche Primzahlen enthalten kann, durch die sie alle teilbar sind, und nicht in höherer Potenz als bei irgend einer von ihnen.

4) Soll von mehreren Zahlen das kleinste gemeinschaftliche Vielfache gefunden werden, so hat man von jeder bei den Zerlegungen vorkommenden Primzahl die höchste Potenz aufzusuchen, in der sie vorkommt, und alle so erhaltenen Primzahlpotenzen zu einem Produkt zu vereinigen.

Denn jedes Vielfache der Zahlen mufs jede in ihnen enthaltene Primzahl und nicht in niederer Potenz enthalten, als sie bei irgend einer der Zahlen vorkommt, also mindestens in so hoher Potenz wie bei den Zahlen mit höchster Potenz.

Als Anwendungen fügen wir die folgenden Bemerkungen hinzu, zu denen der Leser sich selbst die Beweise liefern kann. Die alten Mathematiker nannten eine Zahl vollkommen, wenn sie die Summe ihrer eigentlichen Teiler um 1 übertrifft. Es mag die Bedingung hierfür aufgestellt und der folgende Satz von Euklid bewiesen werden: Ist $2^{n+1}-1$ eine Primzahl, so stellt $2^n(2^{n+1}-1)$ eine vollkommene Zahl dar. Z. B.: $2^2-1=3$, $2^3-1=7$, $2^5-1=31$, $2^7-1=127$ sind Primzahlen, so dafs 6, 28, 496 vollkommene Zahlen sind. Unter befreundeten Zahlen verstand man solche Zahlenpaare, von denen jede die Summe der eigentlichen Teiler der andern um 1 übertrifft. Hiernach ist zu zeigen, dafs 220 und 284, 18416 und 17296, 10744 und 10856, 63020 und 76084, 9437056 und 9363584 solche Zahlen sind.

---

§ 22. Teilbarkeit von Produkten auf einander folgender Zahlen. Fermatscher Satz.

Um eine Zahl hinsichtlich ihrer Teilereigenschaften zu untersuchen, ist nötig zu bestimmen, in welcher höchsten Potenz eine beliebige Primzahl p in ihr enthalten ist. Ist die Zahl ein Produkt mehrerer anderer $a_1, a_2, \ldots a_n$, so läfst sich die Frage darauf zurückführen, wie viele von den Faktoren durch eine Zahl m teilbar sind; wir wollen diese Anzahl kurz durch $\Phi(a_1, a_2, \ldots a_n | m)$ andeuten.

I. Gebrauchen wir das Zeichen $\Psi(a_1, a_2, \ldots a_n | p^k)$ um die Anzahl der Zahlen x der Reihe $a_1, a_2, \ldots a_n$ zu bezeichnen, die durch p^k, aber keine höhere Potenz von p teilbar sind, also der Bedingung

$$(x, p^{k+1}) = p^k$$

genügen, so hat die höchste in dem Produkt $a_1 a_2 \dots a_n$ vorkommende Potenz von p den Exponenten

$$e = \sum_{(k)} k\, \Psi(a_1, a_2, .. a_n \mid p^k)$$

wo die Summation, wie auch bei den folgenden Formeln über alle ganzen positiven Zahlen erstreckt werden kann, aber trotzdem in Wirklichkeit nur über eine endliche Reihe ausgedehnt zu werden braucht, weil keine der Zahlen durch eine beliebig hohe Potenz von p teilbar ist. Nun ist offenbar

$$\Phi(a_1, a_2, \dots, a_n \mid p^i) = \sum_{(k)} \Psi(a_1, a_2, .. a_n \mid p^{i+k})$$

und daher

$$\sum_{(i)} \Phi(a_1, a_2, .. a_n \mid p^i) = \sum_{(i, k)} \Psi(a_1, a_2 .. a_n \mid p^{i+k}).$$

Die Doppelsumme auf der rechten Seite läfst sich aber in der Form einer einfachen Summe

$$\sum_{(h)} h\, \Psi(a_1, a_2, .. a_n \mid p^h)$$

schreiben und stellt also den gesuchten Exponenten e dar, dieser ist also auch gleich

$$e = \sum_{(i)} \Phi(a_1, a_2, \dots a_n \mid p^i).$$

II. Hieraus ergiebt sich folgendes Kriterium: Sind $a_1, a_2 .. a_n$ und $b_1, b_2, \dots b_m$ zwei Reihen von Zahlen, und ist für jede in der letzten vorkommende Primzahlpotenz $m = p^k$

$$\Phi(a_1, a_2, \dots a_n \mid p^k) \geqq \Phi(b_1, b_2, .. b_m \mid p^k),$$

so ist das Produkt $a_1 a_2 .. a_n$ durch das Produkt $b_1 b_2 .. b_m$ teilbar. Denn es enthält das erstere Produkt ja jede im zweiten vorkommende Primzahl in höherer, mindestens aber in gleicher Potenz.

III. Besonders einfach gestaltet sich die Bestimmung der Anzahl Φ bei auf einander folgenden Zahlen. Betrachten wir zunächst den Fall $\Phi(1, 2, \dots a-1, a \mid m)$, so ist sofort erkennbar, dafs sämtliche durch m teilbaren Zahlen dargestellt werden können durch die Reihe

$$m, 2m, \dots . hm,$$

wenn h so gewählt ist, dafs

$$hm \leqq a < (h + 1)m.$$

Setzen wir demgemäfs

$$a = hm + r, \quad \frac{a}{m} = h + \frac{r}{m},$$

so ist $o \leqq r < m$ und h die gröfste in dem Bruche $\frac{a}{m}$ enthaltene ganze Zahl, die wir in der Folge durch $\left[\frac{a}{m}\right]$ bezeichnen wollen. Es ist also

$$\Phi(1, 2, \ldots a-1, a \mid m) = \left[\frac{a}{m}\right].$$

Hiernach ergiebt sich der höchste Exponent, in der die Primzahl p in dem Produkte $1 . 2 \ldots a = a!$ enthalten ist, als gleich

$$e = \sum_i \Phi(1, 2, \ldots a \mid p^i) = \sum_{(i)} \left[\frac{a}{p^i}\right].$$

IV. Sind $a_1, a_2 \ldots a_n$ beliebige ganze Zahlen, so ist offenbar

$$\left[\frac{a_1 + a_2 + \ldots + a_n}{m}\right] \geqq \sum_{(k)} \left[\frac{a_k}{m}\right], \qquad (k = 1, 2, \ldots n)$$

wie man aus den Ungleichungen

$$o \leqq \frac{a_k}{m} - \left[\frac{a_k}{m}\right] < 1 \qquad (k = 1, 2, \ldots n)$$

durch Summation erschliefst, die

$$o \leqq \frac{a_1 + a_2 + \ldots + a_n}{m} - \sum_k \left[\frac{a_k}{m}\right] < n \quad (k = 1, 2, \ldots n)$$

liefert, während andererseits

$$o \leqq \frac{a_1 + a_2 + \ldots + a_n}{m} - \left[\frac{a_1 + a_2 + \ldots + a_n}{m}\right] < 1$$

ist. Bedenken wir nun, dafs

$$\Phi(1, 2, \ldots a_1 + a_2 + \ldots + a_n \mid m) = \left[\frac{a_1 + a_2 + \ldots a_n}{m}\right],$$

$$\Phi(1, 2, \ldots a_i \mid m) \quad = \left[\frac{a_i}{m}\right] \quad (i = 1, 2, \ldots n)$$

ist, so erhalten wir

$$\Phi(1, 2, \ldots a_1 + a_2 + \ldots a_n \mid m)$$
$$\geqq \Phi(1, 2, \ldots a_1, 1, 2 \ldots a_2, \ldots 1, 2 \ldots a_n \mid m).$$

Nach dem in II aufgestellten Kriterium erkennen wir hieraus, dafs $(a_1 + a_2 + \ldots a_n)!$ durch $a_1!\, a_2! \ldots a_n!$ teilbar, oder dafs

$$\frac{(a_1 + a_2 + \ldots + a_n)!}{a_1!\; a_1! \ldots a_n!}$$

eine ganze Zahl ist. Wenn $a_1 + a_2 + \ldots + a_n = p$ eine Primzahl ist, so ergiebt sich noch weiter, dafs $(a_1 + a_2 \ldots + a_n - 1)! = (p-1)!$ durch $a_1!\; a_2!\; a_n!$ teilbar sein mufs, da

$$[(a_1 + a_2 + \ldots a_n)!, a_1!\, a_2! \ldots a_n!] = [p\,(p-1)!, a_1!\, a_2! \ldots a_n!]$$
$$= [p\,(p-1)!, (p-1)!\; a_1!\, a_2! \ldots a_n!, a_1!\, a_2! \ldots a_n!]$$
$$= [(p-1)!\,(p, a_1!\, a_2! \ldots a_n!), a_1!\, a_2! \ldots a_n!]$$
$$= [(p-1)!, a_1!\, a_2! \ldots a_n!]$$

ist.

Dafs die soeben betrachtete Zahl ganz ist, ergiebt sich aus der Theorie der Kombinationen (s. Bd. V dieser Sammlung) auf ganz anderem Wege, denn sie stellt die Anzahl der verschiedenen Arten dar, auf die man $a_1 + a_2 + \ldots + a_n$ Elemente in n Gruppen von je $a_1, a_2, \ldots a_n$ Elementen einteilen kann, und ist so ihrer Natur nach eine ganze Zahl. Solche Zahlen treten auch als Koeffizienten auf beim polynomischen Satz (§ 5), und hierbei ist der oben betrachtete spezielle Fall besonders wichtig. Ist nämlich p eine Primzahl, so sind in der Entwickelung

$$(x_1 + x_2 + \ldots + x_n)^p = \sum_{i_1, i_2, \ldots i_n} \frac{p!}{i_1!\, i_2! \ldots i_n!}\, x_1^{i_1}\, x_2^{i_2} \ldots x_n^{i_n}$$
$$(i_1, i_2, \ldots i_n = 0, 1, \ldots p;\; i_1 + i_2 + \ldots i_n = p)$$

alle Koeffizienten durch p teilbar mit Ausnahme der zu $x_1^p, x_2^p \ldots x_n^p$ gehörigen. Daher ist die Differenz

$$(x_1 + x_2 + \ldots + x_n)^p - (x_1^p + x_2^p + \ldots + x_n^p)$$

durch p teilbar.

Nehmen wir $x_1 = x_2 = \ldots x_n = 1$ an, so folgt, dafs für jede ganze Zahl n stets $n^p - n$ durch p teilbar ist,

wenn p eine Primzahl bedeutet. Wenn n selbst durch p teilbar ist, so ist dieser Satz unmittelbar einleuchtend. Für den Fall aber, dafs n zu p teilerfremd ist, stellt er eine wichtige Beziehung auf. Da dann

$$(n^p - n, p) = [n(n^{p-1} - 1), p] = [(n^{p-1} - 1)(n, p), p] \\ = (n^{p-1} - 1, p)$$

ist, so ergiebt sich der Fermatsche Satz:

Bedeutet p eine Primzahl, n eine nicht durch sie teilbare Zahl, so ist $n^{p-1} - 1$ stets durch p teilbar.

---

§ 23. Die zahlentheoretische Funktion φ (m).

Wir wollen jetzt die Betrachtungen zu Anfang des vorigen Paragraphen nach einer andern Richtung erstrecken. Es seien wieder $a_1, a_2 .. a_m$ eine Reihe von beliebigen Zahlen, und $p_1, p_2, .. p_n$ n vorgelegte Primzahlen. Wir bezeichnen mit $\chi(a_1, a_2, .. a_m \mid p_{i_1} p_{i_2} .. p_{i_k})$ die Anzahl der Zahlen x der ersten Reihe, die durch $p_{i_1}, p_{i_2}, .., p_{i_k}$ aber keine weitere Primzahl teilbar sind, die also der Bedingung

$$(x, p_1 p_2 .. p_n) = p_{i_1} p_{i_2} .. p_{i_k}$$

genügen; $i_1, i_2 .. i_k$ bedeuten hierbei irgend welche k verschiedene Werte aus der Reihe 1, 2, .. n. Dann ergiebt sich als Anzahl der Zahlen $a_1, a_2, .. a_m$, die durch irgend eine der Primzahlen $p_1, p_2 ... p_n$ teilbar sind, der Ausdruck

$$r = \sum_{k} \sum_{(i_1, i_2 .. i_k)} \chi(a_1, a_2, .. a_n \mid p_{i_1} p_{i_2} ... p^{i}_{k}), \qquad (k = 1, 2, ... n)$$

während die übrigen

$$s = m - \sum_{(k)} \sum_{(i_1, i_2 .. i_k)} \chi(a_1, a_2, .. a_m \mid p_{i_1} p_{i_2} .. p_{i_k}) \qquad (k = 1, 2, ... n)$$

Zahlen durch keine der genannten Primzahlen teilbar sind. Die zuerst angegebene Doppelsumme hat 2^n Glieder χ. Nun lassen sich in diesen Summen die χ durch die im vorigen Paragraphen benutzten Anzahlen Φ ersetzen. Es ist nämlich:

$$\Phi(a_1, a_2 \ldots a_m \mid p_{i_1} p_{i_2} \ldots p_{i_k}) = \chi(a_1, a_2 \ldots a_m \mid p_{i_1} p_{i_2} \ldots p_{i_k})$$
$$+ \sum_{(i_{k+1})} \chi(a_1, a_2 \ldots a_m \mid p_{i_1} \ldots p_{i_k} p_{i_{k+1}})$$
$$+ \sum_{(i_{k+1}, i_{k+2})} \chi(a_1, \ldots, a_m \mid p_{i_1} p_{i_2} \ldots p_{i_k} p_{i_{k+1}} p_{i_{k+2}})$$
$$+ \ldots + \chi(a_1, \ldots a_m \mid p_1 p_2 \ldots p_n).$$

Hierbei hat die rechts auftretende Summe

$$\sum_{(i_{k+1}, i_{k+2}, \ldots i_{k+h})} \chi(a_1, a_2 \ldots a_n \mid p_{i_1} \ldots p_{i_k} p_{i_{k+1}} \ldots p_{i_{k+h}}) \quad (h = 1, \ldots n-k)$$

$\frac{(n-k)!}{h!(n-k-h)!}$ Glieder. Durch Summation ergiebt sich nun zunächst

$$\sum_{(i_1, i_2 \ldots i_k)} \Phi(a_1, \ldots, a_m \mid p_{i_1} p_{i_2} \ldots p_{i_k})$$
$$= \sum_h \frac{h!}{k!(h-k)!} \sum_{(i_1 \ldots i_h)} \chi(a_1, \ldots a_m \mid p_{i_1} \ldots p_{i_h})$$
$$(h = k, k+1, \ldots n)$$

und weiter daraus

$$\sum_k (-1)^{k-1} \sum_{(i_1, i_2 \ldots i_k)} \Phi(a_1, a_2 \ldots a_m \mid p_{i_1} \ldots p_{i_k})$$
$$= \sum_{k,h} (-1)^{k-1} \frac{h!}{k!(h-k)!} \sum_{(i_1, i_2 \ldots h)} \chi(a_1, \ldots a_m \mid p_{i_1} \ldots p_{i_h})$$
$$(k = 1 \ldots n;\ h = k, k+1, \ldots n)$$
$$= \sum_{h,k} (-1)^{k-1} \frac{h!}{k!(h-k)!} \sum_{(i_1, i_2 \ldots i_h)} \chi(a_1, \ldots a_m \mid p_{i_1} \ldots p_{i_h})$$
$$(h = 1, \ldots n;\ k = 1, \ldots h).$$

Weil nun aber

$$\sum_k (-1)^{k-1} \frac{h!}{k!(h-k)!} = 1 \qquad (k = 1, \ldots h)$$

ist, so wird die letzte Summe gleich

$$\sum_h \sum_{(i_1 \ldots i_h)} \chi(a_1 \ldots a_m \mid p_{i_1} \ldots p_{i_h}), \qquad (h = 1, \ldots n)$$

also gleich r. Für die oben genannten beiden Anzahlen r und s haben wir also in

$$r = \sum_{k} (-1)^{k-1} \sum_{(i_1 \ldots i_k)} \Phi (a_1 \ldots a_m \mid p_{i_1} \ldots p_{i_k})$$

$$s = m - \sum_{k} (-1)^{k-1} \sum_{(i_1 \ldots i_k)} \Phi (a_1 \ldots a_m \mid p_{i_1} \ldots p_{i_k})$$

$$(k = 1, 2, \ldots n)$$

die gesuchten Darstellungen gefunden.

Nehmen wir jetzt an, dafs $a_1, a_2 \ldots a_m$ die m aufeinander folgenden Zahlen $1, 2, \ldots m$ seien, so ist nach § 22 IV.

$$\Phi (1, 2, \ldots m \mid p_{i_1} \ldots p_{i_k}) = \left[\frac{m}{p_{i_1} p_{i_2} \ldots p_{i_k}}\right].$$

Wenn nun weiter vorausgesetzt wird, dafs m durch $p_1, p_2 \ldots p_n$ teilbar ist, so können wir noch die eckige Klammer weglassen. Es ergiebt sich dann

$$s = m - \sum_{k=1}^{k=n} (-1)^{k-1} \sum_{(i_1 i_2 \ldots i_k)} \frac{m}{p_{i_1} p_{i_2} \ldots p_{i_k}}$$

$$= m\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \ldots \left(1 - \frac{1}{p_n}\right).$$

Nehmen wir nun endlich noch an, dafs m durch keine andern Primzahlen als die genannten teilbar ist, so ist s die Anzahl aller Zahlen unter m, die zu m teilerfremd sind. Bezeichnen wir diese mit φ (m), so ergiebt sich der Satz:

Die Anzahl der unter m gelegenen und zu m teilerfremden Zahlen ist gleich

$$\varphi (m) = m\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right) \ldots \left(1 - \frac{1}{p_n}\right),$$

wenn $p_1, p_2, \ldots p_n$ alle in m aufgehenden Primzahlen bedeuten.

Hiernach ist z. B. $\varphi (2) = 1$, $\varphi (3) = 2$, $\varphi (4) = 2$. Es ist zweckmäfsig $\varphi (1) = 1$ anzunehmen; dies mufs geschehen, wenn man unter φ (m) die Anzahl der Zahlen versteht, die teilerfremd zu m und nicht gröfser als m sind. Wir geben noch folgende Sätze für die Funktion φ an:

1) Ist p eine Primzahl, so ist

$$\varphi (p) = p - 1, \quad \varphi (p^n) = p^{n-1} (p - 1) = p^{n-1} \varphi (p),$$

was man auch sonst leicht ableiten kann.

2) Sind von den Zahlen $m_1, m_2, .. m_n$ immer je zwei zu einander teilerfremd, so ist

$$\varphi(m_1 m_2 .. m_n) = \varphi(m_1)\,\varphi(m_2) \ldots \varphi(m_n),$$

wie leicht daraus erschlossen wird, dafs dann alle Zahlen ohne gemeinschaftliche Primfaktoren sind.

Sehr bemerkenswert ist der folgende Satz:

3) Es ist

$$m = \sum_{(t)} \varphi(t), \qquad (m, t) = t$$

wenn die Summe über sämtliche Teiler t von m erstreckt und $\varphi(1) = 1$ angenommen wird.

Ist nämlich $m = p_1^{e_1}\, p_2^{e_2} .. p_n^{e_n}$, so haben die Teiler die Form (§ 21, 2)

$$t_{i_1 i_2 .. i_n} = p_1^{i_1}\, p_2^{i_2} .. p_n^{i_n},$$
$$(i_1 = 0, 1 \ldots e_1;\ i_2 = 0, 1 \ldots e_2;\ .. i_n = 0, 1 \ldots e_n)$$

und es ist

$$\sum_{(t)} \varphi(t) = \sum_{(i_1, i_2 \ldots i_n)} \varphi(t_{i_1 i_2 .. i_n}) = \sum_{(i_1, i_2 \ldots i_n)} \varphi(p_1^{i_1}\, p_2^{i_2} .. p_n^{i_n})$$

$$= \sum_{(i_1 i_2 .. i_n)} \varphi(p_1^{i_1})\,\varphi(p_2^{i_2}) .. \varphi(p_n^{i_n}) = \prod_{k=1}^{k=n} \sum_{i_k=0}^{i_k=e_k} \varphi(p_k^{i_k})$$

$$= \prod_{k=1}^{k=n} \left[1 + (p-1) \sum_{i_k=0}^{i_k=e_k-1} p_k^{i_k}\right] = \prod_{k=1}^{k=n} p_i^{e_k} = m.$$

Gegenüber dieser Verifikation erscheint dagegen der folgende Beweis ohne jede Rechnung einfacher. Wir teilen alle Zahlen unter m in soviel Gruppen ein, als es Teiler t von m giebt, indem wir alle Zahlen, die mit m den gröfsten gemeinschaftlichen Teiler t haben, in dieselbe Gruppe aufnehmen. Alle diese haben die Form tx, wobei $x < \frac{m}{t}$ und, da aber $(m, tx) = t$ sein soll, $\left(\frac{m}{t}, x\right) = 1$ sein mufs; es gehören demnach zur Gruppe $\varphi\left(\frac{m}{t}\right)$ Zahlen. Summieren wir nun über die Gruppen t, so erhalten wir

$$m = \sum_{(t)} \varphi\left(\frac{m}{t}\right). \qquad (m, t) = t$$

Diese Gleichung ist nur der Form nach von der obigen verschieden, da man ebenso gut die Summation über die Teiler t als über die entsprechenden komplementären Teiler $\frac{m}{t}$ ausführen kann.

---

§ 24. Kongruenz der Zahlen.

Ist die Differenz $a - b$ zweier ganzer Zahlen a und b durch eine dritte Zahl m teilbar, so nennt man die Zahlen a und b kongruent nach m. Nach unseren früheren Bezeichnungen können wir dann schreiben

$$(a - b, m) = m.$$

Aber es ist häufig bequemer, die folgende von Gauſs eingeführte Bezeichnung

$$a \equiv b \mod m$$

anzuwenden, die Kongruenz genannt wird, und bei der dann m der Modul der Kongruenz heiſst. Ist die Differenz $a - b$ nicht durch m teilbar, so nennt man a und b inkongruent nach dem Modul m, in Zeichen

$$a \not\equiv b \mod m.$$

Daſs a durch m teilbar ist, kann man jetzt ausdrücken durch die Kongruenz

$$a \equiv 0 \mod m.$$

Sind zwei Zahlen nach einem Modul kongruent, so sind sie auch nach jedem Teiler desselben kongruent. Sind zwei Zahlen einer dritten kongruent, so sind sie auch unter sich kongruent.

Die Kongruenzen bieten eine groſse Analogie mit den Gleichungen dar, die sich in folgenden Sätzen ausspricht:

Addition und Subtraktion: Aus

$$a \equiv a', \quad b \equiv b' \mod m$$

folgt

$$a + b \equiv a' + b', \quad a - b \equiv a' - b' \mod m.$$

Denn wenn $a - a'$ und $b - b'$ durch den Modul m teilbar sind, so sind es auch nach § 15 die Zahlen

$$(a - a') + (b - b') = (a + b) - (a' + b')$$
$$(a - a') - (b - b') = (a - b) - (a' - b'),$$

woraus die Behauptung sofort hervorgeht.

Als Verallgemeinerung folgt: Ist

$$a_1 \equiv a_1', \quad a_2 \equiv a_2', \quad \ldots \quad a_n \equiv a_n',$$

so ist auch

$$a_1 + a_2 + \ldots + a_n \equiv a_1' + a_2' + \ldots + a_n'.$$

Und weiter, wenn man alle Summanden gleich setzt: Aus

$$a \equiv a'$$

folgt

$$n\,a \equiv n\,a',$$

wo n eine beliebige ganze Zahl ist.

Multiplikation. Aus

$$a \equiv a', \quad b \equiv b'$$

folgt

$$a\,b \equiv a'\,b'.$$

Denn nach der letzten Bemerkung ist

$$a\,b \equiv a'\,b, \quad a'\,b \equiv a'\,b',$$

also ab und $a'b'$ beide kongruent $a'b$, also auch unter sich.

Verallgemeinert man diesen Satz, so ergiebt sich, dafs die Kongruenzen

$$a_1 \equiv a_1', \quad a_2 \equiv a_2', \quad \ldots \quad a_n \equiv a_n'$$

auch die Kongruenz

$$a_1\,a_2 \ldots a_n \equiv a_1'\,a_2' \ldots a_n'$$

nach sich ziehen, und weiter, wenn man die Faktoren gleich setzt, dafs mit

$$a \equiv a'$$

auch

$$a^n \equiv a'^n$$

ist, wenn n eine ganze positive Zahl ist.

Alle diese Sätze kann man nun in einen einzigen Hauptsatz zusammenfassen:

Ist $G(x_1, x_2, .. x_n)$ eine ganze ganzzahlige rationale Funktion der Unbestimmten $x_1, x_2, .. x_n$, so folgt aus

$$a_1 \equiv a_1', \quad a_2 \equiv a_2', \ldots a_n \equiv a_n' \mod m$$

allgemein

$$G(a_1, a_2, .. a_n) \equiv G(a_1', a_2', \ldots a_n') \mod m.$$

Der Beweis ergiebt sich sofort, wenn man bedenkt, dafs eine ganze ganzzahlige rationale Funktion der Gröfsen $x_1, x_2, .. x_n$ aus diesen durch die Operationen der Addition, Subtraktion und Multiplikation erhalten werden kann.

Ganz anders wie zu den Operationen der Addition, Subtraktion und Multiplikation verhalten sich die Kongruenzen in Bezug auf die Division, wo die Analogie mit den Gleichungen aufhört. Denn aus der Kongruenz

$$n a \equiv n a' \mod m$$

kann man nur schliefsen, dafs

$$a \equiv a' \mod \frac{m}{(m, n)}$$

ist. Wir können die Kongruenz nämlich in Form einer Äquivalenz in der Gestalt

$$[n(a - a'), \ m] \sim m$$

schreiben. Wenn wir dann dem Modulsystem auf der linken Seite die Gröfse $m(a - a')$ hinzufügen, so erhalten wir, da

$$[n(a - a'), \ m(a - a')] = (n, m) \cdot (a - a')$$

ist, zunächst

$$[(n, m)(a - a'), m] \sim m$$

und daraus

$$\left[a - a', \frac{m}{(m, n)}\right] \sim \frac{m}{(m, n)},$$

so dafs

$$a \equiv a' \mod \frac{m}{(m, n)}$$

ist.

Anwendungen. Auf die angeführten Sätze lassen sich die bekannten Teilbarkeitskriterien zurückführen, die

man im elementaren Rechnen für dekadische Zahlen anwendet.*) Im dekadischen Zahlensystem wird jede Zahl m durch eine Reihe von Ziffern dargestellt, die die Werte 0, 1, ... 9 annehmen können. Sind $a_0, a_1, \dots a_n$ diese in der Reihe, wie sie von rechts nach links aufeinander folgen, so ist m in der Form darstellbar:

$$m = a_0 + a_1 \,.\, 10 + a_2 \,.\, 10^2 + \dots + a_n \,.\, 10^n,$$

und es ergeben sich folgende Kongruenzen

$$\begin{aligned}
m &\equiv a_0 \mod 10 \\
m &\equiv a_0 + a_1 \,.\, 10 \mod 100 \\
m &\equiv a_0 + a_1 \,.\, 10 + a_1 \,.\, 100 \mod 1000 \\
&\vdots \\
m &\equiv a_0 + a_1 \,.\, 10 + \dots + a_i \,.\, 10^i \mod 10^{i+1}.
\end{aligned}$$

Beachtet man ferner, dafs $10 \equiv 1 \mod 9$, also auch $10^i \equiv 1 \mod 9$ ist, so folgt

$$m \equiv a_0 + a_1 + \dots + a_n \mod 9,$$

wo rechts die sogenannte Quersumme der Zahl auftritt. Ist diese durch 9 teilbar, so ist es die Zahl m also ebenfalls.

Da ferner $10 \equiv -1 \mod 11$ und folglich $10^i \equiv (-1)^i \mod 11$ ist, so hat man

$$m \equiv a_0 - a_1 + a_2 - a_3 + \dots (-1)^n a_n \mod 11.$$

Man kann auf diese Weise für jede Primzahl ein Kriterium aufstellen, doch werden diese nicht mehr so einfach wie die bisherigen. Da z. B. $10 \equiv 3$, $10^2 \equiv 3^2 \equiv 2$, $10^3 \equiv 2 \,.\, 3 \equiv -1$, $10^4 \equiv -3$, $10^5 \equiv -2$, $10^6 \equiv 1 \mod 7$ ist, so erhält man

$$m \equiv a_0 + 3a_1 + 2a_2 - a_4 - 3a_5 - 2a_6 + \dots \mod 7.$$

Dem Leser empfehlen wir, sich weitere solche Kriterien zu bilden, auch dabei an Stelle der Grundzahl 10 eine beliebige andere Zahl g zu setzen. Besonders einfach sind immer die Teilbarkeitskriterien nach den Teilern von g und deren Potenzen, sowie nach $g - 1$ und $g + 1$.

Die entwickelten Sätze über Kongruenzen in Verbindung mit den einfachen Teilbarkeitskriterien können dazu benutzt werden, um die Richtigkeit von Rechnungen zu kontrollieren.

---

*) Vgl. Band I dieser Samml., § 22 (A. d. L.)

Ergiebt sich z. B. aus der Multiplikation zweier Zahlen a und b das Produkt c und sind a', b', c', die kleinsten Reste von a, b, c nach irgend einem Modul m, so mufs $a'b' \equiv c' \bmod m$ sein. Oder wenn sich durch Division von a : b der Quotient g und der Rest c ergiebt, so mufs $a' \equiv g'b' + c' \bmod m$ sein. Die Kontrolle läfst sich bei dekadischen Zahlen leicht ausführen, wenn man $m = 9$ oder 11 annimmt, und man unterscheidet danach eine Neuner-*) oder Elferprobe. Doch ist wohl zu beachten, dafs eine Rechnung falsch sein kann, obgleich sie eine oder mehrere der Proben erfolgreich bestanden hat; denn man braucht ja irgend eine in ihr vorkommende Zahl nur um ein Vielfaches von m zu vermehren oder zu vermindern, um durch die Probe auf keinen Fehler zu stofsen. Erst wenn man die Probe mit allen Primzahlen bis zu der gröfsten durchgeführt hätte, die in den bei den Rechnungen vorkommenden Zahlen enthalten ist, könnte man auf die absolute Richtigkeit einen unanfechtbaren Schlufs ziehen.

---

§ 25. Gruppe des vollständigen Restsystems nach einem Modul.

Jede ganze Zahl kann nach einem ganzzahligen Modul m einer der m Zahlen $0, 1, 2, \ldots m - 1$ kongruent gesetzt oder auf sie reduziert werden. Die Gesamtheit dieser unter sich inkongruenten Zahlen bezeichnet man als ein vollständiges Restsystem nach dem Modul m. Statt der Zahlen $0, 1, .. m - 1$ kann man auch m andere ihnen kongruente, sonst aber beliebig gewählte Zahlen zu Grunde legen, und so überhaupt unendlich viele vollständige Restsysteme nach m bilden, die alle dadurch charakterisiert sind, dafs sie m inkongruente Zahlen enthalten.

Wie nun auch die Zahlen gewählt sein mögen, so ist doch immer die Summe oder Differenz zweier von ihnen einer dritten Zahl des vollständigen Restsystems kongruent. Diese Eigenschaft gilt nicht mehr immer, wenn man einige Zahlen aus dem System entfernt.

---

*) Vgl. Bd. I dieser Samml., S. 121 (A. d. L.).

Wenn jedoch diese Eigenschaft auch einem Teil des Restsystems zukommt, so nennt man diesen eine Gruppe. Es ist klar, dafs die Zahl 0 oder allgemeiner die durch m teilbare Zahl des Restsystems für sich schon eine solche Gruppe bildet, da sich die durch m teilbaren Zahlen durch Addition und Subtraktion reproduzieren. Ferner ergiebt sich, dafs jede Gruppe auch die durch m teilbaren Zahl des Restsystems enthalten mufs, weil die Differenz zweier gleichen Zahlen einerseits in der Gruppe enthalten sein mufs, andererseits aber durch m teilbar ist. Während die durch m teilbare Zahl des Restsystems daher die kleinste Gruppe ist, ist als umfangreichste das volle Restsystem selbst anzusehen.

Wir wollen nun die Konstitution einer solchen Gruppe untersuchen und nehmen dazu an, dafs sie aus den n inkongruenten Zahlen

$$a_0, a_1, \ldots a_{n-1}$$

bestehe. Da der gröfste gemeinschaftliche Teiler

$$a = (a_0, a_1, \ldots a_{n-1}\, m)$$

sich aus $a_0, a_1, \ldots a_{n-1}\, m$ durch wiederholte Addition und Subtraktion herleiten läfst, so mufs für ihn ein Repräsentant $a^{(1)} \equiv a \mod m$ in der Gruppe vorhanden sein; aus a entstehen aber durch fortgesetzte Additionen lauter Vielfache von a, die in der Gruppe ihre Repräsentanten haben müssen. Da aber die Elemente der Gruppe alle Vielfache von a sind, weil sie insgesamt a als gröfsten gemeinschaftlichen Teiler haben, so kann die Gruppe nur aus den n Elementen

$$a^{(1)} \equiv a,\ a^{(2)} \equiv 2a, \ldots a^{(n)} \equiv na \quad \mod m$$

bestehen, und diese müssen inkongruent sein. Eins von ihnen mufs also durch m teilbar sein, das kann aber nur na sein, da sonst nicht sämtliche Vielfache inkongruent wären. Da nun aber a ein Teiler von m ist, so folgt

$$na = m.$$

Die Anzahl n der inkongruenten Elemente der Gruppe oder deren Ordnung und der gröfste gemeinschaftliche Teiler a der Elemente sind also komplementäre Teiler des Moduls m. Ist eine der beiden Zahlen a oder n gegeben, so ist die Gruppe vollständig bestimmt, bis auf die Re-

präsentanten, deren Wahl auf unendlich viele Weisen abgeändert werden kann. Es giebt so viele verschiedene Gruppen, wie es Teiler der Zahl m giebt. Ist der Modul eine Primzahl, so giebt es also aufser dem Element Null und dem vollständigen Restsystem keine Gruppen mehr.

Die Zahl a, durch deren wiederholte Addition die Gruppe von der Ordnung n gebildet werden kann, nennen wir ein Grundelement der Gruppe, und zwar ist sie das kleinste Grundelement, das es giebt, da sie ein Teiler von m ist. Jedes andere Grundelement b hat mit m das kleinste Grundelement a als gröfsten gemeinschaftlichen Teiler, so dafs also

$$(m, b) = a$$

ist. Denn das kleinste Vielfache von b, das durch m teilbar ist, ist $\frac{mb}{(m,b)}$; wenn aber dieses der Zahl $nb = \frac{mb}{a}$ gleich sein soll, so mufs $(m, b) = a$ sein. Solche Zahlen b existieren aber in der Anzahl $\varphi\left(\frac{m}{a}\right)$. Es giebt demnach $\varphi(n)$ inkongruente Grundelemente einer Gruppe nter Ordnung.

---

§ 26. Gruppe des verkürzten Restsystems nach einem Modul.

Die in § 23 betrachteten inkongruenten Reste, die zum Modul m teilerfremd und in der Anzahl $\varphi(m)$ vorhanden sind, bezeichnen wir als das verkürzte Restsystem nach dem Modul m, wobei wir jedoch hier statt der früher in Betracht gezogenen kleinsten Reste auch $\varphi(m)$ ihnen kongruente, sonst aber beliebig gewählte Repräsentanten zulassen wollen, die alle inkongruent und teilerfremd zu m sind. Sind a und b zwei Zahlen in diesem Restsystem, so ist ihr Produkt ab teilerfremd zu m, weil

$$(ab, m) = (ab, am, bm, m^2) = (a, m)(b, m) = 1$$

ist, und mufs daher einer Zahl c des Systems kongruent sein, da dieses ja alle inkongruenten teilerfremden Zahlen zu m enthält. Es kann nun vorkommen, dafs einem Teil des verkürzten Restsystems die Eigenschaft zukommt, dafs

das Produkt zweier in ihm enthaltenen Reste immer wieder im Teilsystem durch einen mod m kongruenten Repräsentanten vertreten ist. Einen solchen Teil nennen wir eine Gruppe. Die Einheit bildet für sich allein schon eine Gruppe, da die wiederholte Multiplikation sie nicht ändert, und es ist auch leicht zu zeigen, dafs sie die kleinste Gruppe darstellt, da jede Gruppe das Element 1 enthalten mufs. Ist nämlich a ein Element einer Gruppe, so sind die Potenzen von a gleichfalls Elemente. Solche kann man aber in unendlicher Menge bilden; da nun die Gruppe an inkongruenten Zahlen nur eine beschränkte Anzahl enthalten kann, so kann man (auf unendlich mannigfache Art) zwei kongruente Potenzen a^i und a^k der Gruppe ermitteln. Ist $i > k$, so folgt aber aus $a^i \equiv a^k \bmod m$, da $(a, m) = 1$ ist, $a^{i-k} \equiv 1 \bmod m$, und a^{i-k} ist offenbar auch in der Gruppe enthalten.

Das soeben angewandte Beweisverfahren liefert zugleich ein einfaches Mittel, um eine Gruppe mit Hilfe eines einzigen Elementes a herzustellen. Bilden wir nämlich von diesem Element die Reihe der Potenzen, und nennen wir n die kleinste Zahl, für die $a^n \equiv 1 \bmod m$ ist, so sind die n Elemente

$$1, a, a^2, \ldots a^{n-2}, a^{n-1}$$

inkongruent, denn wäre $a^i \equiv a^k$, $n > i > k$, so müfste $a^{i-k} \equiv 1$, $i - k < n$ sein, was der Annahme über n widerspricht. Die hingeschriebenen Elemente bilden aber auch eine Gruppe, denn es ist

$$a^i a^k \equiv a^l \bmod m,$$

wenn

$$i + k \equiv l \bmod n$$

gesetzt wird, wo l mit einer der Zahlen $0, 1, \ldots n - 1$ übereinstimmend gewählt werden kann. Setzt man die Potenzbildung fort, so wiederholen sich die Elemente periodisch; man erkennt z. B., dafs

$$a^n \equiv 1, \quad a^{n+1} \equiv a, \ldots a^{2n-1} \equiv a^{n-1}$$

u. s. w. ist. Die Zahl n, die Ordnung der definierten Gruppe, nennt man auch die Ordnung, Periode oder den Exponenten von a in Bezug auf den Modul m; a selbst, sowie jedes andere Element, durch dessen Poten-

zierung die Gruppe gebildet werden kann, das Grundelement der Gruppe. Doch lassen sich nicht alle Gruppen durch ein solches Grundelement erzeugen; für den Fall, daſs der Modul eine Primzahl ist, ist das allerdings der Fall, wie wir später erkennen werden (§ 71). Aus der angestellten Betrachtung erhellt, daſs jede Gruppe stets ein Element $\overline{a}$ enthält, das, mit einem anderen Element a multipliziert, das Einheitselement erzeugt und das inverse Element zu a genannt wird. Hat a in Bezug auf m die Periode n, so ist

$$a\, a^{n-1} \equiv a^{n-1}\, a \equiv 1 \bmod m,$$

also a^{n-1} das inverse Element zu a, so daſs $\overline{a} \equiv a^{n-1}$ mod m ist.

Wir wollen nun annehmen, daſs zwei Gruppen G und H vorliegen, von denen die erste alle Elemente der zweiten enthält. H sei von der Ordnung r, d. h. enthalte r inkongruente Elemente

$$a_0 \equiv 1,\ a_1, a_2, \ldots a_{r-1}$$

und keine weiteren. Multiplizieren wir alle Elemente mit einer zu m teilerfremden Zahl b, so erhalten wir eine Reihe

$$a_0 b \equiv b,\ a_1 b, a_2 b, \ldots a_{r-1} b$$

von lauter inkongruenten und zu m teilerfremden Zahlen, die entweder alle Elemente oder kein einziges von H enthält, je nachdem entweder b in H enthalten ist oder nicht. Ist nämlich b in H enthalten, so ist es auch jedes Element $a_i b$ $(i = 0, 1 \ldots r-1)$, und umgekehrt ist $a_i b$ in H enthalten, so auch $\overline{a_i}\, a_i\, b \equiv b$. Betrachten wir allgemeiner zwei Reihen

$$a_0 b' \equiv b',\ a_1 b',\ a_2 b', \ldots a_{r-1} b'$$

und

$$a_0 b'' \equiv b'',\ a_1 b'',\ a_2 b'', \ldots a_{r-1} b'',$$

die aus der Multiplikation der Elemente von H mit zwei zu m teilerfremden Zahlen b' und b'' entstehen, so haben diese entweder alle Elemente oder kein einziges gemeinschaftlich, je nachdem $b'\overline{b''}$ (oder auch $b''\overline{b'}$) in H enthalten ist oder nicht. Ist nämlich $b'\overline{b''}$ in H enthalten, so gilt dasselbe von $a_i b' \overline{b''}$, so daſs jedes Element $a_i b'$ der ersten Reihe, weil es kongruent $(a_i b' \overline{b''})\, b''$ ist, auch in der zweiten

Reihe vorkommt. Ebenso ergiebt sich dann, dafs auch jedes Element der zweiten Reihe in der ersten vertreten ist, so dafs beide dieselben Elemente enthalten. Haben aber die beiden Reihen auch nur ein einziges Element gemeinsam, etwa $a_i b' \equiv a_k b''$, so folgt aus $b' \overline{b''} = \overline{a_i} a_k$, dafs $b' \overline{b''}$ in H enthalten ist.

Wenn nun die Gruppe H nicht alle Elemente von G enthält, so kann man aus dieser ein Element b_1 bestimmen, so dafs die Reihe

$$a_0 b_1, \; a_0 b_2, \ldots a_{r-1} b_1,$$

r neue in G, aber nicht in H vertretene Elemente liefert. Sind hiermit noch nicht alle Elemente von G aufgestellt, so kann man in G ein neues Element b_2 finden, so dafs die Reihe

$$a_0 b_2, \; a_1 b_2, \ldots a_{r-1} b_2$$

lauter neue Elemente liefert. So kann man fortfahren und die Gruppe G erschöpfen, die ja eine beschränkte Zahl von Elementen hat; also mufs man zuletzt zu einem Element b_{s-1} gelangen, so dafs durch die Reihe

$$a_0 b_{s-1}, \; a_1 b_{s-1}, \ldots \; a_{r-1} b_{s-1}$$

im Verein mit den vorhergehenden alle Elemente von G zur Darstellung gelangen. Ist n die Ordnung von G, so folgt $n = rs$ und damit der Satz:

Enthält eine Gruppe alle Elemente einer andern Gruppe, so ist ihre Ordnung ein Vielfaches der Ordnung dieser andern Gruppe.

Da alle Gruppen in der umfassendsten Gruppe von der Ordnung $\varphi(m)$ enthalten sind, so folgt:

Die Ordnungszahl jeder Gruppe von teilerfremden Zahlen des Moduls m ist ein Teiler von $\varphi(m)$.

Nun haben wir aber gesehen, dafs sich aus jeder zu m teilerfremden Zahl a durch Potenzieren eine Gruppe von der Ordnung n bilden läfst, wenn n den Exponenten von a in Bezug auf m darstellt, d. h. die kleinste Zahl, für die

$$a^n \equiv 1 \bmod m$$

ist. Da $\varphi(m)$ ein Vielfaches von n ist, so gilt um so mehr die Kongruenz

$$a^{\varphi(m)} \equiv 1 \mod m$$

für jede zum Modul m teilerfremde Zahl a. Diese Kongruenz nennt man den verallgemeinerten Fermatschen Satz. Ist nämlich $m = p$ eine Primzahl, so ist $\varphi(p) = p - 1$ und daher

$$a^{p-1} \equiv 1 \mod p,$$

wenn a nicht durch p teilbar ist. Diese Kongruenz drückt aber den Fermatschen Satz aus, den wir in § 22 durch ganz andere Betrachtungen gefunden haben.

IV. Abschnitt.

Lineare Kongruenzen mit einer Unbekannten.

---

§ 27. Reduktion linearer Modulsysteme auf eine einfachere Form.

Während wir im vorigen Abschnitt nur Modulsysteme in Betracht gezogen haben, deren Elemente bestimmte ganze Zahlen sind, wollen wir uns jetzt auch zu solchen wenden, in denen Unbestimmte vorkommen, dabei allerdings uns zunächst auf den einfachsten Fall beschränken, dafs die Elemente ganze lineare Funktionen mit einer einzigen unbestimmten Gröfse x sind.

Es seien $a_1 x + b_1$, $a_2 x + b_2$ zwei solche Funktionen, deren Koeffizienten a_1, b_1, a_2, b_2 ganze Zahlen sind. Wir betrachten das Modulsystem

$$(a_1 x + b_1, \; a_2 x + b_2)$$

und wollen zunächst zeigen, dafs sich dieses stets auf die Form

$$(ax + b, c)$$

reduzieren läfst. Sind a_1 und a_2 von Null verschieden und ist $a_1 \geqq a_2$, was ohne Beschränkung der Allgemeinheit vorausgesetzt werden kann, so können wir a_1 in ein Vielfaches $g a_2$ von a_2 und einen Rest a_3, der kleiner als a_2 ist, zerlegen, so dafs also

$$a_1 = g a_2 + a_3, \quad a_3 < a_2$$

ist. Setzen wir nun zur Abkürzung

$$b_1 = g b_2 + b_3,$$

so ist

$$a_1 x + b_1 = g(a_2 x + b_2) + a_3 x + b_3,$$

und daraus folgt

$$(a_1 x + b_1,\ a_2 x + b_2) = (a_2 x + b_2,\ a_3 x + b_3).$$

Zugleich ist aber

$$(a_1, a_2) = (a_2, a_3), \quad (b_1, b_2) = (b_2, b_3)$$

und, wie man durch einfache Rechnung erkennt,

$$a_1 b_2 - a_2 b_1 = a_2 b_3 - a_3 b_2.$$

Nennen wir die Koeffizientenverbindung $a_1 b_2 - a_2 b_1$ die Determinante des Modulsystems $(a_1 x + b_1,\ a_2 x + b_2)$, so ist $a_2 b_3 - a_3 b_3$ die des transformierten Systems $(a_2 x + b_2,\ a_3 x + b_3)$, und wir können den Inhalt der letzten Gleichung einfach so formulieren, dafs die Determinante durch vorhin angegebene Transformation keine Änderung erleidet.

Ist nun $a_3 \neq 0$, so kann man auf das System

$$(a_2 x + b_2,\ a_3 x + b_3)$$

dasselbe Verfahren anwenden wie auf das ursprüngliche, und kann so fortfahren, bis man auf ein Modulsystem kommt, in dem der Koeffizient von x in der einen Funktion verschwindet. Das mufs geschehen, weil die Koeffizienten $a_1, a_2, a_3, \ldots$ an Gröfse beständig abnehmen und genau dieselbe Zahlenreihe darstellen, wie sie bei der Bestimmung des gröfsten gemeinschaftlichen Teilers von a_1 und a_2 auftritt. Bezeichnen wir das entstehende Modulsystem mit

$$(ax + b,\ c),$$

so ist infolge der obigen Bemerkungen

$$(a_1, a_2) = a, \quad (b_1, b_2) = (b, c), \quad ac = a_1 b_2 - a_2 b_1,$$

da die Determinante des zuletzt erhaltenen Systems durch das Produkt ac dargestellt wird. Wir erhalten somit folgenden Satz:

Das Modulsystem

$$(a_1 x + b_1,\quad a_2 x + b_2)$$

läfst sich immer auf die Form bringen

$$(ax + b, c),$$

wo

$$a = (a_1, a_2), \ (b, c) = (b_1, b_2), \ c = \frac{a_1 b_2 - a_2 b_1}{(a_1, a_2)}$$

ist.

Es ist nicht schwer, diesen Satz auf eine beliebige Anzahl von linearen Funktionen zu verallgemeinern. Wir überlassen dem Leser die nähere Ausführung und geben nur das Resultat an, dafs man jedes Modulsystem von der Form

$$(a_1 x + b_1, \ a_2 x + b_2, \ \ldots a_n x + b_n)$$

auf die Form

$$(ax + b, c)$$

bringen kann, wo

$$a = (a_1, a_2, \ldots a_n), \ (b, c) = (b_1, b_2, \ldots b_n)$$
$$ac = (a_1 b_2 - a_2 b_1, \ a_1 b_3 - a_3 b_1, \ \ldots a_{n-1} b_n - a_n b_{n-1})$$

ist.

---

§ 28. Weitere Reduktion des Systems $(ax + b, m)$.

Die Form, auf die wir die im vorigen Paragraphen betrachteten Modulsysteme gebracht haben, ist nun noch einer weiteren Reduktion fähig, wenn wir — was allerdings von vornherein stillschweigend angenommen wurde, da sonst der Begriff des Modulsystems, wie wir ihn im vorigen Abschnitt kennen gelernt haben, keinen Sinn mehr hätte und erweitert werden müfste — hinzunehmen, dafs die Unbestimmte x eine ganze Zahl bedeuten soll.

Wir können dann dem System

$$(ax + b, c)$$

als ein beliebiges Vielfache von c die Gröfse cx hinzufügen, wobei dann aus dem System

$$(ax + b, cx, c)$$

das Teilsystem

$$(ax + b, cx)$$

nach der Methode des vorigen Paragraphen weiter behandelt werden kann. Dieses läfst sich nämlich reduzieren auf

$$\left(a'x + b', \frac{bc}{a'}\right),$$

wo

$$a' = (a, c)$$

gesetzt ist. Fügen wir dem Teilsystem wieder die Zahl c hinzu und beachten, dafs

$$\left(\frac{bc}{a'}, c\right) = \frac{(a, b, c)\,c}{(a, c)}$$

ist, so erhalten wir an Stelle des ursprünglichen Modulsystems das äquivalente von ähnlicher Form

$$(a'x + b', c'),$$

wo

$$a' = (a, c), \qquad c' = \frac{(a, b, c)\,c}{(a, c)}$$

ist. Wir können nun annehmen, dafs $a \neq 0$ ist, da sonst das ursprüngliche System einfach den Wert (b, c) hätte. Dann ist aber

$$a' \leqq a, \qquad c' \leqq c.$$

Wenn $a' = a$, $c' = c$ wird, oder anders ausgedrückt, wenn

$$a = (a, c) = (a, b, c)$$

ist, so hat die Transformation keine Änderung des Systems $(ax + b, c)$ zur Folge. Wir nennen es dann r e d u z i e r t. Ist es also nicht reduziert, so mufs sich bei jeder Transformation wenigstens eine der beiden Zahlen a oder c verkleinern. Da aber eine solche Verkleinerung nicht beliebig fortgesetzt werden kann, weil es nur eine beschränkte Menge von ganzen Zahlen giebt, die kleiner als a resp. c sind, so mufs man durch genügende Wiederholung der Transformation zu einem reduzierten System gelangen. Um die Form des Systems etwas genauer bestimmen zu können, beachten wir, dafs

$$(a, b, c) = (a', b', c')$$

ist. Nennen wir (a, b, c) den w e s e n t l i c h e n T e i l e r des M o d u l s y s t e m s und bezeichnen ihn mit d, so zeigt sich, dafs er eine bei allen Transformationen unveränderliche

Gröfse ist. Hieraus ergiebt sich, dafs das Modulsystem $(ax + b, c)$ auf die Form

$$d(x + r, m)$$

gebracht werden kann, wo m ein Teiler von $\frac{c}{d}$ ist. In dem besonderen Falle, wo $m = 1$ wird, kann man noch die Gröfse x hinzufügen; da aber $(x + r, x) = (r, x)$, also $(x + r, x, 1) = (r, x, 1) = 1$ wird, so ist dann das System seinem wesentlichen Teiler äquivalent. Die Gröfse m kann noch genauer bestimmt werden, doch wollen wir uns hierauf nicht einlassen, da dies für das Folgende nicht von Wichtigkeit ist.

Wir fügen nun noch einige Beispiele zur Veranschaulichung bei:

1) $(3x + 5, 17) = (3x + 5, 17, 17x)$
$= [3x + 5, 17, 17x - 5(3x + 5) + 17]$
$= (3x + 5, 17, 2x - 8) = [3x + 5 - (2x - 8), 17, 2x - 8]$
$= (x + 13, 17, 2x - 8) = [x + 13, 17, 2x - 8 - 2(x + 13)]$
$= (x + 13, 17, 34) = (x + 13, 17) = (x - 4, 17).$

2) $(10x + 35, 65) = 5(2x + 7, 13) = 5(2x + 7, 13, 13x)$
$= 5[2x + 7, 13, 13x - 6(2x + 7) + 3 . 13]$
$= 5(2x + 7, 13, x - 3)$
$= 5[2x + 7 - 2(x - 3), 13, x - 3] = 5(13, x - 3)$
$= 5(x - 3, 13).$

3) $(2x + 3, 8) = (2x + 3, 8, 8x)$
$= [2x + 3, 8, 8x - 4(2x + 3) + 8]$
$= (2x + 3, 8, 4) = (2x + 3, 4) = (2x + 3, 4, 4x)$
$= [2x + 3, 4, 4x - 2(2x + 3) + 4] = (2x + 3, 4, 2)$
$= (2x + 3, 2) = (2x + 3, 2, 2x) = (2x + 3 - 2x, 2, 2x)$
$= (3, 2, 2x) = (3, 2) = 1.$

Für die folgenden Beispiele mag der Leser die Reduktionen selbst ausführen:

4) $(6x + 5, 8) = 1,$
5) $(6x + 5, 14) = (x + 2, 7),$
6) $(330x + 13, 840) = (x - 1, 7).$

§ 29. Auflösung linearer Kongruenzen mit einer Unbekannten. Diophantische Gleichungen.

Mit Hilfe der entwickelten Reduktionsmethode ist es nun möglich zu entscheiden, ob ein gegebenes lineares Modulsystem $(ax + b, c)$ irgend einem gegebenen Wert n äquivalent werden kann, wenn man der unbestimmten Zahl x geeignete Werte beilegt. Legen wir das reduzierte System $d(x + r, m)$ zu Grunde, so ergiebt sich aus

$$d(x + r, m) = n$$

erstens, dafs n durch d, und zweitens, wenn man unter dieser Voraussetzung

$$(x + r, m) = \frac{n}{d}$$

ableitet, dafs m durch $\frac{n}{d}$ teilbar sein mufs. Diese beiden Bedingungen sind aber nicht nur notwendig, sondern auch hinreichend für die Lösbarkeit der Aufgabe. Betrachten wir die letzte Äquivalenz, so sagt diese aus, dafs $x + r$ und m den gröfsten gemeinschaftlichen Teiler $\frac{n}{d}$ haben. Solche Zahlen $x + r$ giebt es unendlich viele, denn aus einer einzigen solchen Zahl kann man durch Hinzufügung von Vielfachen von m beliebig viele andere ableiten. Wir brauchen uns daher nur auf solche zu beschränken, die kleiner als m sind. Die Anzahl dieser ist nach § 23 gleich $\varphi\left(\frac{md}{n}\right)$. Nennen wir eine von ihnen ϱ, so ist $x + r \equiv \varrho \mod m$, also

$$x \equiv \varrho - r \mod m$$

ein der Aufgabe genügender Wert von x.

Ein ganz spezieller Fall dieser Aufgabe ist die Lösung der Kongruenz ersten Grades mit einer Unbekannten

$$ax + b \equiv 0 \mod m,$$

die als Äquivalenz geschrieben werden kann in der Gestalt

$$(ax + b, m) = m.$$

Es sei $d(x+\varrho, \mu)$ das reduzierte System zu $(ax+b, m)$ so ist nach der Bemerkung des vorigen Paragraphen

$$d = (a, b, m), \left(\frac{m}{d}, \mu\right) = \mu,$$

während die in diesem abgeleitete erste Bedingung von selbst erfüllt ist, die zweite aber die Form

$$\left(\mu, \frac{m}{d}\right) = \frac{m}{d}$$

annimmt. Hieraus ergiebt sich sofort, dafs

$$\mu = \frac{m}{d}$$

sein mufs. Die Bedingung hierfür läfst sich aber leicht aufstellen. Unterwirft man nämlich das Modulsystem $(ax+b, m)$ einer einmaligen Transformation, so erhält man das äquivalente

$$(a'x+b', m'),$$

wo

$$a' = (a, m), \quad m' = \frac{md}{a'}$$

ist. Da nun $d\mu = m$ nicht gröfser als $\frac{md}{a'}$ sein kann und sicher nicht kleiner ist, weil $a' \geq d$ ist, so mufs $d = a'$ sein. Dann wird aber auch sicher das System schon bei der ersten Transformation reduziert auf die gewünschte Form. Hieraus ergiebt sich der Satz:

Die hinreichende und notwendige Bedingung für die Lösbarkeit der Kongruenz

$$ax+b \equiv 0 \bmod m$$

lautet

$$(a, b, m) = (a, m).$$

Übrigens läfst sich diese Bedingung auf einem weit einfacheren Wege ableiten, wenn man einerseits beachtet, dafs die Kongruenz aussagt, dafs sich b linear und homogen durch a und m ausdrücken lassen soll und daher durch den gröfsten gemeinschaftlichen Teiler (a, m) teilbar sein mufs, andererseits die Bedingung bedeutet, dafs sich b linear und

homogen durch a und m darstellen läfst, dafs also auch umgekehrt die Kongruenz besteht.

Beachten wir ferner, dafs aus dem reduzierten System

$$\left(x+\varrho, \frac{m}{d}\right)=\frac{m}{d}$$

folgt, so ergiebt sich, dafs die Kongruenz nur ein einziges Lösungssystem von der Form

$$x \equiv -\varrho \bmod \frac{m}{d}$$

nach dem Modul $\frac{m}{d}$ hat.

Häufig verlangt man aber alle Lösungsysteme nicht nach dem Modul $\frac{m}{d}$, sondern nach dem Modul m der Kongruenz zu wissen. Nach dem Modul m hat die Kongruenz dann d Lösungsysteme, denn man kann einer bestimmten Lösung ja Vielfache von $\frac{m}{d}$ hinzufügen; solche Vielfache, die kleiner als m sind, giebt es aber im ganzen d, nämlich das 0, 1, .. (d — 1) fache. Ist also x_0 eine Lösung, so ergeben sich aus ihr die weiteren

$$x_0+\frac{m}{d},\quad x_0+2\frac{m}{d},\ \ldots\ x_0+(d-1)\frac{m}{d},$$

die alle nach m inkongruent sind.

Die linearen Kongruenzen werden oft eingekleidet in die Form von unbestimmten Gleichungen von der Gestalt

$$ax+by=c,$$

für die Lösungen in ganzen Zahlen gesucht werden. Eine solche Gleichung nennt man eine Diophantische Gleichung, weil zuerst Diophantos von Alexandria solche durch Raten gelöst hat. Die oben hingeschriebene ist offenbar gleichwertig mit einer der beiden Kongruenzen

$$ax-c\equiv 0 \bmod b,\quad by-c\equiv 0 \bmod a.$$

Die hinreichende und notwendige Bedingung für die Auflösbarkeit lautet also

$$(a, b, c)=(a, b).$$

Wir fügen nun noch einige Beispiele hinzu:

1)
$$7x + 5 \equiv 0 \mod 17$$
$$(7x + 5, 17) = (x + 8, 17)$$
$$x \equiv 9 \mod 17.$$

2)
$$25x + 111y + 13 = 0$$
$$(25x + 13, 111) = (x - 35, 111)$$
$$x \equiv 35 \mod 111.$$

Setzt man demgemäfs, unter t eine beliebige ganze Zahl verstehend,
$$x = 35 + 111t,$$
so ergiebt sich durch Einsetzen in die Diophantische Gleichung der zugehörige Wert von y in der Form
$$y = -8 - 25t.$$

3)
$$30x + 18 \equiv 0 \mod 138$$
$$(30x + 18, 138) = 6\,(5x + 3, 23) = 6\,(x - 4, 23).$$

Daher folgt
$$6\,(x - 4, 23) = 138, \quad (x - 4, 23) = 23, x \equiv 4 \mod 23.$$

Die Lösungsysteme nach 138 sind daher
$$x \equiv 4, 27, 50, 73, 96, 119 \mod 178.$$

4) Soll die Äquivalenz
$$(93x + 27, 72) = 12$$
gelöst werden, so findet man
$$(93x + 27, 72) = (21x + 27, 72) = 3\,(x - 9, 24),$$
es mufs also sein
$$3\,(x - 9, 24) = 12, \quad (x - 9, 24) = 4.$$

Die $\varphi\left(\frac{24}{4}\right) = 2$ Lösungen nach 24 sind $x - 9 \equiv 4$, 20 mod 24, also
$$x \equiv 5, 13 \mod 24,$$
oder wir erhalten 2 . 3 Lösungssysteme nach 72
$$x \equiv 5, 29, 53;\ 13, 37, 61 \mod 72.$$

§ 30. Systeme von linearen Kongruenzen.

Sind mehrere lineare Kongruenzen nach demselben Modul m zu lösen

$$a_i x + b_i \equiv 0 \mod m, \qquad (i = 1, 2, \ldots n)$$

so kann man diese in die Form einer einzigen Äquivalenz

$$(a_1 x + b_1, a_2 x + b_2, \ldots a_n x + b_n, m) = m$$

zusammenfassen. Auf das lineare Modulsystem der linken Seite kann man dann die in § 27 und 28 beschriebene Reduktionsmethode anwenden. Wir wollen dies hier aber nicht ausführen, sondern es der Übung des Lesers anheimgeben, zu beweisen, dafs die Bedingungen der Lösbarkeit sich in der Form:

$$(a_i b_k - a_k b_i, a_1 m, a_2 m, \ldots a_n m, b_1 m, b_2 m, \ldots b_n m, m^2)$$
$$= (a_1, a_2, \ldots a_n) m \qquad (i, k = 1, 2, \ldots n)$$

darstellen lassen, und dafs dann ein einziges Lösungssystem nach dem Modul $\dfrac{m}{(a_1, a_2 \ldots a_n, m)}$ existiert.

Sind die Moduln der einzelnen linearen Kongrenzen verschieden, so kann man die Kongruenzen auch durch eine einzige Äquivalenz ersetzen, und es bieten sich dann weitere Schwierigkeiten nicht mehr dar. Sind die Kongruenzen in der Form

$$a_i x + b_i \equiv 0 \mod m_i \qquad (i = 1, 2, \ldots n)$$

gegeben und versteht man unter m irgend ein Vielfaches der sämtlichen Moduln $m_1, m_2, \ldots m_n$, also z. B. das kleinste gemeinschaftliche Vielfache, oder, wenn man lieber will, das Produkt der Moduln, und setzt demgemäfs

$$m = l_i m_i,$$

so ist das System von Kongruenzen völlig gleichbedeutend mit der Äquivalenz

$$[l_1 (a_1 x + b_1), l_2 (a_2 x + b_2), \ldots l_n (a_n x + b_n), m] = m.$$

Denn multipliziert man die $i$te Kongruenz mit l_i, so erhält man die neue

$$l_i (a_i x + b_i) \equiv 0 \mod m$$

und kann aus dieser auch wieder rückwärts die ursprüngliche ableiten, da $(m, l_i) = l_i$ ist. Die Zusammenfassung aller so erhaltenen Kongruenzen mit demselben Modul führt aber auf die hingeschriebene Äquivalenz. Die notwendige und hinreichende Bedingung für die Lösbarkeit läfst sich schreiben in der Gestalt

$$[(a_i b_k - a_k b_i) l_i l_k, (a_1 l_1, a_2 l_2 \ldots a_n l_n, b_1 l_1 \ldots b_n l_n) m, m^2] = (a_1 l_1, a_2 l_2, \ldots a_n l_n) m,$$

und es ergiebt sich dann ein einziges Lösungssystem nach dem Modul $\frac{m}{(a_1 l_1, \ldots a_n l_n, m)}$.

---

§ 31. Besondere Systeme von linearen Kongruenzen.

Obgleich durch die einfachen Betrachtungen des letzten Paragraphen das Problem der Lösung linearer Kongruenzen zum völligen Abschlufs gebracht ist, wollen wir doch nicht verabsäumen, auf gewisse vorher anzubringende Vereinfachungen hinzuweisen. Man kann nämlich jede Kongruenz des Systems auch zunächst für sich behandeln, ohne auf die übrigen Rücksicht zu nehmen. Vorausgesetzt nun, dafs alle einzelnen Kongruenzen sich als lösbar erweisen, läfst sich das Problem auf die Lösung von Kongruenzen der Form

$$x \equiv a_i \mod m_i \qquad (i = 1, 2, \ldots n)$$

zurückführen und in folgender einfachen Weise formulieren: Es sollen alle Zahlen x bestimmt werden, die in Bezug auf n gegebene Moduln $m_1, m_2, \ldots m_n$, n gegebene Reste $a_1, a_2, \ldots a_n$ haben.

Wir wollen nun annehmen, dafs immer je zwei dieser Moduln zu einander teilerfremd sind, so dafs

$$(m_i, m_k) = 1 \qquad (i \neq k)$$

sein soll, um hieran einige wichtige Folgerungen zu knüpfen und nachher zu zeigen, dafs sich das Problem immer so einrichten läfst, dafs diese Voraussetzung erfüllt ist. Setzen wir

$$m = m_1 m_2 \ldots m_n, \quad m = l_i m_i,$$

so läfst sich, da

$$(l_1, l_2, .. l_n) = 1, \ [l_i\, l_k\, (a_i - a_k),\ m] = m \qquad (i \neq k)$$

ist, die linke Seite der Äquivalenz

$$[l_1\, (x - a_1),\ l_2\, (x - a_2), \ldots l_n\, (x - a_n),\ m] = m$$

nach § 27 auf die Form

$$(x - a,\ m)$$

transformieren, und daraus geht hervor, dafs es immer und zwar nur einziges Lösungssystem nach dem Modul m giebt, das den Kongruenzen Genüge leistet. Bei der Reduktion des Systems spielen nun allein die Koeffizienten l_1, l_2 .. l_n von x eine Rolle (§ 27), und daraus ergiebt sich, dafs bei der Darstellung von a in der Form

$$a \equiv \sum_i h_i\, a_i \mod m \qquad (i = 1, 2, \ldots n)$$

die Koeffizienten h_i nur allein von ihnen oder von m_1, m_2 ...m_n abhängig sind. Nehmen wir $a_1 = 0, .. a_{i-1} = 0,\ a_i = 1,$ $a_{i+1} = 0, .. a_n = 0$ an, so ergiebt sich durch

$$[l_1\, x,\ l_2\, x, .. l_{i-1}\, x,\ l_i\, (x - 1),\ l_{i+1}\, x, .. l_n\, x,\ m] = (x - h_i,\ m)$$

eine Definition von h_i mod m. Man kann nämlich h_i, wie hieraus unmittelbar folgt, wenn man $x = h_i$ setzt, durch die Kongruenzen:

$$h_i \equiv \delta_{ik} \mod m_k \qquad (i, k = 1, 2 .. n)$$

definieren, wo das Zeichen δ_{ik} die Zahl 1 oder 0 bedeutet, je nachdem i und k einander gleich sind oder verschiedene Werte haben. Es ist also h_i ein Vielfaches von l_i, und wenn man $h_i = l_i\, x_i$ setzt, so genügt x_i der Kongruenz

$$l_i\, x_i \equiv 1 \mod m_i. \qquad (i = 1, 2, \ldots n)$$

Endlich kann man die Gröfsen x_i als solche auffassen, die der Kongruenz

$$\sum_i l_i\, x_i \equiv 1 \mod m \qquad (i = 1, 2, \ldots n)$$

genügen, wie solche wegen $(l_1, l_2, .. l_n, m) = 1$ existieren müssen. Dann ergiebt sich aus

$$x - a \equiv \sum_i l_i\, x_i\, (x - a_i) \mod m \qquad (i = 1, 2 \ldots n)$$

sofort:

$$a \equiv \sum_i l_i\, x_i\, a_i \equiv \sum_i h_i\, a_i \mod m. \qquad (i = 1, 2 \ldots n).$$

Da die Kongruenzen

$$x \equiv a_i \bmod m_i \qquad (i = 1, 2 \ldots n)$$

bei gegebenen Werten von a_1 mod m_1, a_2 mod $m_2, \ldots a_n$ mod m_n immer ein einziges Lösungssystem a mod m haben, und umgekehrt einem solchen immer nur ein einziges Lösungssystem a_1 mod m_1, a_2 mod $m_2, \ldots a_n$ mod m_n entspricht, so durchläuft a ein vollständiges Restsystem nach m, wenn $a_1, a_2, \ldots a_n$ resp. solche nach $m_1, m_2, \ldots m_n$ durchlaufen und umgekehrt. Aus der Äquivalenz (§ 20)

$$(a, m_1)\,(a, m_2) \ldots (a, m_n) = (a, m)$$

ergiebt sich nun:

$$(a_1, m_1)\,(a_2, m_2) \ldots (a_n, m_n) = (a, m),$$

da $a \equiv a_i \bmod m_i$, also $(a, m_i) = (a_i, m_i)$ ist. Ist also $(a, m) = 1$, so ergiebt sich $(a_1, m_1) = 1, \ldots (a_n, m_n) = 1$ und umgekehrt. Hieraus folgt, daſs a das verkürzte Restsystem nach m durchläuft, wenn $a_1, a_2, \ldots a_n$ resp. eben solche nach $m_1, m_2, \ldots m_n$ durchlaufen und umgekehrt. Daher ist

$$\varphi(m) = \varphi(m_1)\,\varphi(m_2) \ldots \varphi(m_n),$$

wie wir auf anderem Wege in § 23 gefunden hatten. Von dieser Formel ausgehend kann man leicht den allgemeinen Ausdruck für $\varphi(m)$ ableiten, wenn man m in lauter Primzahlpotenzen zerlegt und beachtet, daſs

$$\varphi(p) = p - 1, \quad \varphi(p^n) = p^n(p - 1)$$

für jede Primzahl p ist, was sich unabhängig von der Entwickelung in § 23 leicht zeigen läſst.

Aus der oben abgeleiteten Kongruenz

$$a \equiv \sum_i l_i x_i a_i \bmod m \qquad (i = 1, 2, \ldots n)$$

folgt, daſs der Bruch $\frac{a}{m}$ sich in der Form

$$\frac{a}{m} = \sum_i \frac{a_i x_i}{m_i} + g \qquad (i = 1, 2, \ldots n)$$

darstellen läſst, wo g eine ganze Zahl bedeutet. Setzen wir $a_i x_i = y_i$, so ist y_i definiert die Kongruenz

$$l_i y_i \equiv a \bmod m_i. \qquad (i = 1, 2, \ldots n)$$

Man kann daher y_i auf eine Weise als eine positive Zahl unter m_i so bestimmen, dafs durch die Zerlegung

$$\frac{a}{m} = \sum_i \frac{y_i}{m_i}$$

jeder echte Bruch $\frac{a}{m}$ in eine Summe von echten Partialbrüchen dargestellt wird, die lauter Nenner besitzen, die gegen einander teilerfremd sind.

Lassen wir jetzt die zu Anfang unserer Untersuchung aufgestellte Annahme, dafs immer je zwei der Moduln m_i zu einander teilerfremd sind, fallen, so können wir jeden Modul in ein Produkt von Zahlen zerlegen, die gegenseitig zu einander teilerfremd sind, z. B. dadurch, dafs man als solche die im Modul enthaltenen höchsten Primzahlpotenzen annimmt, und dann aus jeder Kongruenz soviel neue Kongruenzen ableitet, als Faktoren vorhanden sind. Wie wir soeben bewiesen haben, sind diese einzelnen abgeleiteten der einzigen ursprünglichen gleichwertig. Verfährt man so mit allen Kongruenzen, so hat man nur noch diejenigen, bei denen die Moduln Potenzen derselben Primzahl sind, daraufhin zu untersuchen, ob sie sich auch widersprechen, und, wenn das nicht der Fall ist, nur die mit den höchsten Primzahlpotenzen in Betracht zu ziehen. Somit kommt die Untersuchung auf den schon behandelten Fall zurück, dafs immer je zwei Moduln zu einander teilerfremd sind.

Beispiele.

1)

$$x \equiv a \mod 7$$
$$x \equiv b \mod 5$$
$$x \equiv c \mod 3$$

$$[15(x - a),\ 21(x - 6),\ 35(x - c),\ 105]$$
$$= (x - 15a - 21b + 35c,\ 105)$$
$$x \equiv 15a + 21b - 35c \mod 105.$$

2)

$$x \equiv a \mod 8$$
$$x \equiv b \mod 9$$
$$x \equiv c \mod 7$$

$$[72(x - a),\ 56(x - 6),\ 72(x - c),\ 504]$$
$$= (x + 63a + 224b + 216c,\ 504)$$
$$x \equiv -63a - 224b - 216c \mod 504.$$

3) Die Kongruenz

$$23x \equiv 667 \mod 693,$$

kann, da $693 = 3^2 . 7 . 11$ ist, in drei andere zerlegt werden, wobei sich

$$x \equiv 2 \mod 9, \quad x \equiv 1 \mod 7, \quad x \equiv 7 \mod 11$$

ergiebt. Das allgemeine System

$$x \equiv a \mod 9, \quad x \equiv b \mod 7, \quad x \equiv c \mod 11$$

hat die Lösung

$$x \equiv 154a + 99b + 252c \mod 693,$$

so dafs in unserem Falle $x \equiv 29 \mod 693$ wird.

4) Das System

$$\begin{aligned} x &\equiv 6 \mod 18 \\ x &\equiv 15 \mod 21 \\ x &\equiv 12 \mod 24 \\ x &\equiv 3 \mod 33 \end{aligned}$$

zerlegt sich in die folgenden Kongruenzen

$$\begin{aligned} x &\equiv 0 \mod 2, & x &\equiv 6 \mod 9 \\ x &\equiv 0 \mod 3, & x &\equiv 1 \mod 7 \\ x &\equiv 4 \mod 8, & x &\equiv 0 \mod 3 \\ x &\equiv 0 \mod 3, & x &\equiv 3 \mod 11, \end{aligned}$$

die keinen Widerspruch enthalten, so dafs man jetzt das System

$$\begin{aligned} x &\equiv a \mod 8 \\ x &\equiv b \mod 9 \\ x &\equiv c \mod 7 \\ x &\equiv d \mod 11 \end{aligned}$$

zu lösen hat, wo a, b, c, d resp. gleich 4, 6, 1, 3 sind; dessen allgemeine Lösung

$$x \equiv -2079a - 1232b + 792c + 2520d \mod 5544,$$

also für unsern speziellen Fall $x \equiv 3732 \mod 5544$ ergiebt.

V. Abschnitt.

Permutationsgruppen.

§ 32. Permutationen.

Wie wir später sehen werden, hat man in der Algebra häufig rationale Funktionen zu betrachten, die ihren Wert oder ihre Form nicht ändern, wenn man die Variabeln gewissen Vertauschungen unterwirft. Es ist daher zweckmäſsig, die hierbei auftretenden Möglichkeiten einer genaueren Untersuchung zu unterziehen.

Sind $a_1, a_2, \ldots a_n$ n verschiedene Elemente, so kann man die Anzahl der verschiedenen Anordnungen, die offenbar nur von der Zahl n abhängt und demgemäſs mit $P_{(n)}$ bezeichnet werden kann, auf folgende Weise bestimmen: Wir denken uns alle möglichen Anordnungen so in Abteilungen geordnet, daſs wir alle mit demselben Elemente beginnenden derselben Abteilung einverleiben. So entstehen n Abteilungen $A_1, A_2, \ldots A_n$, so daſs A_1 alle mit a_1, A_2 alle mit a_2, $\ldots A_n$ alle mit a_n beginnenden Anordnungen enthält. Die Anzahl der Anordnungen in jeder dieser Abteilungen ist aber gleich groſs, nämlich gleich der Anzahl $P_{(n-1)}$ der möglichen Anordnungen der auf das erste Element folgenden $(n-1)$ Elemente. So entsteht die Gleichung

$$P_{(n)} = n\, P_{(n-1)}.$$

Ersetzen wir hierin n durch $n-1$, $n-2$, $\ldots 2$, so erhalten wir noch hinzu

$$P_{(n-1)} = (n-1)\, P_{(n-2)}$$
$$\vdots$$
$$P_{(2)} = 2 P_{(1)},$$

und es ergiebt sich dann durch Multiplikation, wenn man noch bedenkt, dafs $P(1) = 1$ ist,

$$P(n) = 2 . 3 . . . n.$$

Das Produkt $1 . 2 . 3 . . n$ bezeichnet man häufig abgekürzt durch $n!$ oder $\Pi(n)$.

Die Anzahl der verschiedenen Anordnungen von n Elementen ist also gleich n!

Den Übergang von einer Anordnung zu einer anderen nennt man eine Permutation.*). Sind $a_1, a_2, \ldots a_n$ die n Elemente, die durch $a_{i_1}, a_{i_2}, \ldots a_{i_n}$ ersetzt werden sollen, wo $i_1, i_2, \ldots i_n$ n lauter verschiedene Zahlen aus der Reihe 1, 2, . . n sind, so bezeichnet man die hierdurch ausgedrückte Permutation kurz durch

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ a_{i_1} & a_{i_2} & & a_{i_n} \end{pmatrix}.$$

Die Elemente der oberen Reihe können hierbei in ganz willkürlicher Reihenfolge hingeschrieben werden; hat man aber eine Folge ausgewählt, so sind die der unteren Reihe völlig bestimmt. Wenn man auch noch in dem Falle, wo die Elemente keiner neuen Anordnung unterworfen werden, von einer (identischen) Permutation spricht, so giebt es überhaupt n! Permutationen von n Elementen.

---

§ 33. Zerlegung einer Permutation in Cyklen.

Ist eine beliebige Permutation gegeben, so kann man aus ihr ein Element a_0 herausgreifen, das durch sie in a_1 übergeführt werden mag, darauf das Element a_2 bestimmen, in das a_1 vermöge der Permutation übergeht, und so fortfahren, bis man auf ein Element a_{n-1} stöfst, das durch die Permutation wieder in das ursprüngliche Element a_0 übergeht. Zu einem solchen mufs man stets gelangen, weil die Permutation nur eine endliche Zahl von zu vertauschenden Elementen enthält. Die so bestimmte Permutation

$$A = \begin{pmatrix} a_0 & a_1 & \ldots & a_{n-2} & a_{n-1} \\ a_1 & a_2 & & a_{n-1} & a_0 \end{pmatrix},$$

*) Vgl. auch „Permut.“ in Bd. V dieser Samml. (A. d. L.)

deren Elemente sich so anordnen lassen, dafs jedes Element in das folgende, das letzte aber in das erste übergeführt wird, nennt man eine cyklische Permutation oder kurz einen Cyklus. Einen Cyklus von zwei Elementen nennt man auch eine Transposition.

Einen Cyklus bezeichnet man gewöhnlich kurz dadurch, dafs man die Elemente so nebeneinander schreibt, von links nach rechts, wie sie in einander übergeführt werden, die obige also in der Form

$$A = (a_0, a_1, a_2, \ldots a_{n-1}),$$

doch steht hierbei vor dem ersten Element a_0 nicht das Element a_{n-1}, hinter dem letzten nicht a_0. Hierin liegt eine Willkürlichkeit, die offenbar vermieden würde, wenn man die Elemente nicht in einer geraden Linie, sondern in einer geschlossenen, etwa einem Kreise, anordnen wollte, was aus leicht erklärlichen praktischen Gründen nicht geschieht. Teilt man einen Kreis in n gleiche Teile, ordnet den Teilpunkten der Reihe nach die Elemente a_0, $a_1 \ldots a_{n-1}$ zu und läfst ihn dann den nten Teil einer vollen Umdrehung vollführen, so dafs der Teilpunkt a_0 auf den Teilpunkt a_1 fällt, so fällt a_1 auf a_2, a_2 auf a_3, u. s. w., a_{n-2} auf a_{n-1} und schliefslich a_{n-1} auf a_0. Daraus erklärt sich die Bezeichnung cyklische Permutation. Eine Folge der oben erwähnten Willkürlichkeit in der Bezeichnung ist es, dafs man einen Cyklus auf soviel Arten bezeichnen kann, als er Elemente enthält, da jedes Element zu Anfang hingeschrieben werden kann.

Wie wir oben gesehen hatten, kann man aus jeder Permutation einen Cyklus aussondern. Enthält die Permutation nun noch andere Elemente, wie sie in dem ausgesonderten Cyklus vorkommen, so kann man ein neues Element b_0 herausgreifen, das wieder einen Cyklus B nach sich zieht. So kann man, wenn die Elemente noch nicht erschöpft worden, fortfahren, mufs aber wegen der endlichen Anzahl der Elemente zu einem Abschlufs gelangen. Dann hat man die Permutation P in lauter Cyklen A, B, C... ohne gemeinsame Elemente zerlegt. Um dies anzudeuten, schreiben wir symbolisch

$$P = A\,B\,C \ldots$$

Hierbei werden alle Elemente weggelassen, die nicht geändert werden.

Die Zerlegung einer Permutation in Cyklen ist eine völlig eindeutige bis auf die Anordnung der Cyklen, die vollständig willkürlich ist.

---

§ 34. Zusammensetzung von Permutationen.

Sind zwei Permutationen, die wir kurz durch A und B bezeichnen, derselben Elemente gegeben, so kann man immer eine dritte Permutation angeben, die denselben Übergang hervorbringt, wie wenn man zuerst die Permutation A, darauf die Permutation B ausführt. Man bezeichnet diese neue Permutation durch AB und nennt sie die Resultante von A und B, A und B die Komponenten und zwar A die erste und B die zweite Komponente. Die Bildung von AB bezeichnet man als Zusammensetzung der Permutationen A und B. Schon im vorigen Paragraphen haben wir diese Bezeichnung angewandt, wobei dann allerdings die zusammenzusetzenden Permutationen keine Elemente gemeinsam haben. Offenbar kann man die Cyklen A, B, C... am Schlusse des vorigen Paragraphen auch als Permutationen aller Elemente ansehen, bei denen die nicht in ihnen vorkommenden Elemente durch sich selbst zu ersetzen sind und daher nicht hingeschrieben zu werden brauchen.

Die Zusammensetzung kann sich auch auf mehrere Permutationen beziehen. Nehmen wir den Fall an, dafs drei Permutationen A, B, C vorliegen, so können wir die Resultante ABC auf zwei Arten bilden. Soll erst A mit B und darauf AB mit C zusammengesetzt werden, so schreiben wir die Resultante in der Form (AB) C. Wird dagegen erst BC gebildet und erfolgt darauf die Zusammensetzung von A mit BC, so wenden wir die Bezeichnung A (BC) an. Diese Unterscheidung, die begrifflich notwendig erscheint, ist aber, wenn man nur auf das Resultat Gewicht legt, unnötig, denn es gilt stets das associative Gesetz

$$(AB)\,C = A\,(BC).$$

Wird nämlich durch die Permutation A ein beliebiges Element a in b, durch B das Element b in c und durch C endlich c in d übergeführt, so ist leicht zu erkennen, dafs a schliefslich stets durch d ersetzt wird, gleichviel ob die Zusammensetzung nach der Formel (AB) C oder A (BC) erfolgt. Denn nach der ersteren wird erst a in b, darauf in c und endlich in d übergeführt, während nach der zweiten zuerst a in b, darauf b in c, dann aber in d und somit a in d übergeführt wird. Wenn wir nur auf das Resultat sehen, so können wir die Resultante der drei Permutationen kurz durch ABC bezeichnen, eine Klammer nur etwa dann gebrauchen, um damit anzudeuten, dafs eine der beiden Zusammensetzungen zu einer besonders einfachen Rechnung führt.

---

§ 35. Inverse Permutation. Potenzen und Ordnung einer Permutation.

Da die Bezeichnung der Komposition der Permutationen völlig mit der der Multiplikation übereinstimmt, so bezeichnet man die identische Permutation häufig durch das Zeichen 1; wie die Zahl 1 bei der Multiplikation keine Änderung hervorbringt, so auch die identische Permutation bei der Zusammensetzung mit irgend einer Permutation. Einen Anstofs kann eine solche Bezeichnung nicht erregen. Vielfach hat man sogar die Zusammensetzung der Permutationen direkt als Multiplikation bezeichnet, obgleich ihr die Eigenschaft der Kommutativität abgeht.

Zu jeder Permutation P kann man stets eine solche $\bar{P}$ bilden, die mit ihr komponiert, keine Vertauschung herbeiführt, so dafs

$$P\bar{P} = 1, \qquad \bar{P}P = 1$$

ist. Führt dann P das Element a in b über, so mufs $\bar{P}$ b in a überführen. Man nennt die Permutation $\bar{P}$ die inverse Permutation von P.

Um die wiederholte Zusammensetzung einer Permutation P mit sich selbst kurz darzustellen, wendet man dieselbe Bezeichnungsweise wie bei den Potenzen an; man nennt auch

die entstehenden Permutationen P^2, P^3, u. s. w. meistens Potenzen statt Iterationen von P. Man setzt also:

$$PP = P^2, \qquad P^2 P = P^3, \qquad P^3 P = P^4 \ldots,$$

und allgemein gilt als Definition, wenn n eine beliebige ganze Zahl bedeutet,

$$P^n P = P^{n+1}.$$

Es gelten auch hier, wenn a und b zwei ganze Zahlen darstellen, die Gesetze

$$P^a P^b = P^{a+b},$$
$$(P^a)^b = P^{ab} = (P^b)^a, \quad P^a \overline{P}^a = 1,$$

von denen sich das zweite und dritte leicht aus dem ersten ableiten läfst. Das erste läfst sich aber aus der Associativität der Zusammensetzung der Permutationen durch vollständige Induktion beweisen. Ist nämlich das Gesetz für die Zahlen a und b gültig, so ergiebt sich aus

$$P^a P^{b+1} = P^a P^b P = P^{a+b} P = P^{a+b+1},$$

dafs es für jede gröfsere Zahl b, und dann aus

$$P^{a+1} P^b = P^a P P^b = P^a P^{b+1} = P^{a+b+1},$$

dafs es auch für jede gröfsere Zahl a gilt. Für $a = 1$, $b = 1$ ist das Gesetz aber richtig, folglich gilt es allgemein.

Fährt man in der Bildung der Potenzen einer Permutation fort, so mufs man, da die Anzahl der möglichen Permutationen eine endliche ist, während die Potenzen beliebig fortgesetzt werden können, auf unendlich viele Weisen verschiedene ganze Zahl i und k erhalten, so dafs

$$P^i = P^k$$

ist. Komponiert man nun unter der Voraussetzung $i > k$ auf beiden Seiten mit $\overline{P}^k$, so ergiebt sich

$$P^i \overline{P}^k = P^{i-k} (P^k \overline{P}^k) = P^{i-k} = 1.$$

Es giebt also auch eine kleinste Zahl n, für die $P^n = 1$ ist. Dann sind die Potenzen

$$1, P, P^2, P^3, \ldots P^{n-1}$$

von einander verschieden, und es können durch weitere Zusammensetzung keine andern Permutationen aus P erzeugt

werden. Allgemein ist $P^i = P_k$, wenn $i \equiv k \bmod n$ ist. Die Zahl n nennt man dann die Ordnung der Permutation P.

Die inverse Permutation $\overline{P}$ ist gleich P^{n-1}, und man kann hierfür auch P^{-1} schreiben, indem man für die Exponenten auch negative Werte zuläſst.

---

§ 36. Beispiele und Anwendungen.

1) Zerlege in Cyklen:

$$\begin{pmatrix}1&2&3&4&5&6&7&8\\3&4&5&8&1&7&2&6\end{pmatrix}, \begin{pmatrix}1&2&3&4&5&6&7&8&9\\9&5&6&2&1&7&4&3&8\end{pmatrix},$$

$$\begin{pmatrix}a&b&c&d&e&f&g&h&i&k&l\\h&k&d&f&b&i&l&g&e&c&a\end{pmatrix}, \begin{pmatrix}a&b&c&d&e&f&g&h&i&k\\f&d&g&k&c&i&h&e&a&b\end{pmatrix},$$

$$\begin{pmatrix}x_1&x_2&x_3&x_4&x_5&x_6\\x_1&x_4&x_5&x_6&x_3&x_2\end{pmatrix}.$$

2) Bilde von den obigen Permutationen die inversen und die Potenzen. Zerlege diese dann in Cyklen.

3) Wie groſs ist die Ordnung des n-gliedrigen Cyklus

$$(a_0, a_1, a_2, \dots a_{n-1})?$$

4) Beweise, daſs

$$(a, a_1, a_2, \dots a_n) = (a, a_1)(a_1, a_2, \dots a_n)$$

und ebenso

$$(a_1, a_2, \dots a_n) = (a_1, a_2)(a_1, a_3, \dots a_n)$$

$$\vdots$$

$$(a, a_{n-1}, a_n) = (a, a_{n-1})(a, a_n)$$

ist und leite daraus ab, daſs

$$(a, a_1, a_2, \dots a_n) = (a, a_1)(a_1, a_2) \dots (a, a_n)$$

ist.

5) Beweise, daſs sich jede Permutation aus Cyklen von 2 Elementen (mit einem gemeinsamen Elemente) darstellen läſst. (S. u. § 40.)

6) Beweise, daſs

$$(a, a_1 \dots a_n)(a, b) = (a, a_1, \dots a_n, b)$$

$$(a, a_1 \dots a_n) = (a, a_1, \dots a_n, b)(a, b)$$

ist.

7) Beweise, dafs ferner

$$(a, a_1, \ldots a_{n-1}, b, b_1 \ldots b_{m-1})(a, b)$$
$$= (a, a_1 \ldots a_{n-1})(b, b_1 \ldots b_{m-1})$$

ist, ebenso dafs

$$(a, a_1 \ldots a_{n-1})(b, b_1 \ldots b_{m-1})(a, b)$$
$$= (a, a_1, \ldots a_{n-1}, b, b_1 \ldots b_{m-1})$$

ist. Wie kann man das zweite Resultat unmittelbar aus dem ersten ableiten?

8) Beweise, dafs

$$\begin{pmatrix} a_0' & a_1' & \ldots & a'_{n-1} \\ a_0 & a_1 & \ldots & a_{n-1} \end{pmatrix} (a_0, a_1, \ldots a_{n-1}) \begin{pmatrix} a_0 & a_1 & \ldots & a_{n-1} \\ a_0' & a_1' & \ldots & a'_{n-1} \end{pmatrix}$$
$$= (a_0', a_1', a_2' \ldots a'_{n-1}),$$

und dafs bei

$$P = (a_0, a_1, \ldots a_{n-1})(b_0, b_1 \ldots b_{m-1})$$
$$Q = \begin{pmatrix} a_0 & a_1, & \ldots & a_{n-1} & b_0 & b_1 & \ldots & b_{m-1} \\ a_0' & a_1', & \ldots & a'_{n-1} & b_0' & b_1' & \ldots & b'_{m-1} \end{pmatrix}$$

die Beziehung

$$\bar{Q} P Q = (a_0' a_1', \ldots a'_{n-1})(b_0' b_1' \ldots b'_{m-1})$$

gilt. (S. u. § 40.)

9) Hat eine Permutation P die Ordnung n, so hat P^a die Ordnung $\frac{n}{(a, n)}$.

10) Untersuche die in 9) gestellte Frage unter der Voraussetzung, dafs P ein n-gliedriger Cyklus ist. Vgl. 3).

11) Wie viele verschiedene Potenzen von P giebt es dann, die die Ordnung m haben, wenn m ein Teiler von n ist?

12) Welches ist die Ordnungszahl einer Permutation P, die sich in n Cyklen ohne gemeinsame Elemente $A_1, A_2, \ldots A_n$ zerlegen läfst, die resp. $a_1, a_2, \ldots a_n$ Elemente enthalten?

13) Beweise, dafs wenn $P = A_1 A_2 \ldots A_n$ die Zerlegung der Permutation P in Cyklen $A_1, A_2 \ldots A_n$ darstellt, allgemein $P^m = A_1^m A_2^m \ldots A_n^m$ ist.

---

§ 37. Permutationsgruppen.

Das einfachste Beispiel für die Anwendung der Permutationen bietet sich bei Betrachtung der Formveränderung rationaler Funktionen mehrerer Veränderlichen $x_1, x_2 \ldots x_n$ dar unter der Voraussetzung, dafs für diese letzteren ganz beliebige Werte gesetzt, also auch Permutationen mit ihnen vorgenommen werden dürfen. Ist eine solche Funktion F gleich Null, so verschwindet sie identisch, und man darf aus $F = 0$ dann schliefsen $F_P = 0$, wobei F_P die Funktion bedeutet, die aus F durch Anwendung einer beliebigen Permutation P der Variabeln entsteht. Ist allgemein $F = G$, so ist auch $F_P = G_P$. Wendet man auf F zuerst die Permutation A an, darauf auf F_A die Permutation B, so entsteht F_{AB}, das aus F durch Anwendung von AB sofort hervorgeht.

Wird nun eine Funktion F nicht geändert bei Anwendung zweier Permutationen A und B, ist also

$$F_A = F, \quad F_B = F,$$

so folgt hieraus mit Anwendung der soeben dargelegten Bemerkung:

$$F_{AB} = F_B, \quad F_{BA} = F_A,$$

es ist also auch

$$F_{AB} = F_{BA} = F.$$

Wird also eine Funktion bei Anwendung zweier Permutationen nicht geändert, so bringt auch deren Zusammensetzung keine Änderung hervor. Die Gesamtheit aller Permutationen, die eine Funktion nicht ändern, hat also die Eigenschaft, dafs die Zusammensetzung zweier wieder zu ihr gehörige ergiebt. Ein solches System nennt man eine Permutationsgruppe, die Anzahl der zu ihr gehörigen Permutationen deren Ordnung. Das System aller möglichen Permutationen von n Elementen ist offenbar die umfassendste Gruppe, die sich bilden läfst. Eine Funktion, die sich durch deren Permutationen nicht ändert, heifst eine symmetrische Funktion. Um aus irgendwelchen gegebenen Permutationen eine Gruppe zu bilden, hat man in der Zusammensetzung je zweier so lange fortzufahren, bis keine neuen Permutationen mehr entstehen. Der einfachste Fall ist der,

dafs nur eine einzige Permutation gegeben ist, wobei die einzelnen Potenzen zu bilden sind, und die Ordnung der Gruppe mit der der Permutation übereinstimmt. Hieraus folgt, dafs in jeder Gruppe die identische Permutation und zu jeder Permutation die inverse vorhanden ist.

Wir wollen nun annehmen, dafs eine Gruppe G alle Permutationen einer Gruppe H enthält, die wir durch

$$A_0, A_1, \ldots A_{n-1}$$

bezeichnen. Nehmen wir irgend eine Permutation B aus G und bilden die Reihe der Permutationen

$$A_0 B, A_1 B, \ldots A_{n-1} B,$$

so sind alle diese von einander verschieden, wenn $A_0, A_1 \ldots A_{n-1}$ verschieden sind, da aus $A_i B = A_k B$ durch Komposition mit B^{-1} sofort $A_i = A_k$ folgen würde. Die neue Reihe enthält nun entweder alle Elemente von G oder kein einziges, je nachdem B in G enthalten ist oder nicht. Und betrachten wir zwei Reihen

$$A_0 B', A_1 B', \ldots\ldots A_{n-1} B',$$
$$A_0 B'', A_1 B'', \ldots A_{n-1} B'',$$

so ergiebt sich als hinreichende und notwendige Bedingung dafür, dafs sie alle, resp. keine Elemente gemeinschaftlich haben, je nachdem $B'\overline{B''}$ (oder auch $B''\overline{B'}$) in G enthalten ist oder nicht. Dieselben Betrachtungen, die wir in § 26 zur Anwendung gebracht haben, lassen sich ohne irgend eine Modifikation auf den vorliegenden Fall übertragen, und wir gewinnen auf diese Weise den Satz:

Enthält eine Gruppe G alle Permutationen einer andern Gruppe H, so ist ihre Ordnung ein Vielfaches der Ordnung dieser letzteren Gruppe H.

Da die Gesamtheit der Permutationen von n Elementen eine Gruppe von der Ordnung n! darstellt, so mufs die Ordnung irgend einer Gruppe von diesen Permutationen ein Teiler von n! sein. Es ist also auch die Ordnung jeder Permutation in n! als Teiler enthalten.

---

§ 38. Konjugierte und ausgezeichnete Gruppen.

Alle Permutationen einer Gruppe G, die eine gegebene Funktion F ungeändert lassen, bilden eine Gruppe H, die die Permutationen

$$A_0, A_1, \ldots A_{n-1}$$

enthalten möge, so dafs alle andern die Form oder den Wert von F ändern. Man kann aber, wie sich leicht einsehen läfst, eine Anzahl Permutationen $B_1, B_2, \ldots B_{r-1}$ aus G bestimmen, die alle möglichen Werte von F hervorbringen, die mit Hülfe sämtlicher Permutationen von G zu erhalten sind. Wir können sie durch

$$F, F_{B_1}, F_{B_2}, \ldots F_{B_{r-1}}$$

darstellen; sie werden die zu einander konjugierten Funktionen in G genannt. Die Anzahl r ist ein Teiler von n! Zu den einzelnen konjugierten Funktionen gehören nun Permutationsgruppen, die auch zu einander konjugiert in Bezug auf G genannt werden, und zu deren Bestimmung wir jetzt übergehen.

Läfst die Permutation A die Funktion F_B ungeändert, so ist

$$F_{BA} = F_B.$$

Hieraus folgt mit Anwendung der Permutation B^{-1}

$$F_{BAB^{-1}} = F,$$

dafs also dann F durch die Permutation BAB^{-1} nicht geändert wird, die daher in H enthalten sein und in der Form

$$BAB^{-1} = A_i$$

darstellbar sein mufs. Hieraus folgt aber

$$A = B^{-1}A_iB.$$

Alle so möglichen Permutationen

$$A_0, B^{-1}A_1B, B^{-1}A_2B, \ldots B^{-1}A_{n-1}B$$

lassen in der That F_B ungeändert. Sie bilden ferner auch wirklich eine Gruppe; denn durch Zusammensetzung zweier von ihnen $B^{-1}A_iB$, $B^{-1}A_kB$ ergiebt sich $B^{-1}A_iA_kB$, oder weil $A_iA_k = A_l$ gesetzt werden kann, wo A_l in G enthalten ist, $B^{-1}A_lB$, was wieder unter den hingeschriebenen

Permutationen vorkommt. Alle konjugierten Gruppen in G haben also dieselbe Ordnung. Wir erhalten sie, wenn wir für B aufser der identischen Permutation die oben bezeichneten Permutationen $B_1, B_2, \ldots B_{r-1}$ setzen, in der Darstellung

$$H = \{A_0, A_1, \ldots A_{n-1}\}$$
$$H_1 = \{B_1^{-1} A_0 B_1, B_1^{-1} A_1 B_1, \ldots B_1^{-1} A_{n-1} B_1\}$$
$$\vdots$$
$$H_{r-1} = \{B_{r-1}^{-1} A_0 B_{r-1}, B_{r-1}^{-1} A_1 B_{r-1}, \ldots B_{r-1}^{-1} A_{n-1} B_{r-1}\}.$$

Doch brauchen diese r konjugierten Gruppen nicht verschieden zu sein, und es ist der Fall besonders wichtig, dafs sie alle identisch sind. Dies trifft z. B. zu, wenn alle Elemente in G gegen einander vertauschbar sind; denn aus $A_i B = B A_i$ folgt sofort $B^{-1} A_i B = A_i$. Deswegen haben wir auch in §§ 25 und 26 nicht nötig gehabt, auf den Begriff der konjugierten Gruppen einzugehen, weil diese alle als identisch ausgefallen wären. Sind alle zu H in Bezug auf G konjugierten Gruppen identisch, so nennt man H eine ausgezeichnete Gruppe oder Normalgruppe von G. Solche Gruppen spielen in der Algebra, wie wir später sehen werden, eine hervorragende Rolle. Eine Gruppe, die keine Normalgruppe enthält, nennt man eine einfache Gruppe. Bei allen andern Grupppen, die zusammengesetzte genannt werden, entsteht die Aufgabe, sie zu zerlegen, d. h. eine Normalgruppe in ihr von möglichst hoher Ordnung zu bestimmen, von dieser dann wieder eine Normalgruppe zu ermitteln, bis man zuletzt auf eine einfache Gruppe kommt, in der nur noch das Einheitselement als ausgezeichnetes zu betrachten ist. Eine Gruppe, deren Ordnung eine Primzahl ist, ist stets eine einfache Gruppe, weil sie aufser der identischen überhaupt keine Untergruppen enthält. In § 11 werden wir Beispiele von einfachen Gruppen nennen lernen, deren Ordnungen zusammengesetzte Zahlen sind, und die in der Algebra eine wichtige Bedeutung haben.

Wenn man die konjugierten Permutationsgruppen, deren Elemente allgemein die Form $\overline{Q} P Q$ haben, bilden will, so ist die folgende Bemerkung von Nutzen. Es sei P in Cyklen zerlegt ohne gemeinschaftliche Elemente (§ 32) in der Form

$$P = (a_0, a_1, \ldots a_{n-1})\,(b_0, b_1, \ldots b_{m-1})\,(c_0, c_1, \ldots c_{l-1}) \ldots$$

und

$$Q = \begin{pmatrix} a_0\ a_1 \ldots a_{n-1}\ b_0\ b_1 \ldots b_{m-1}\ c_0\ c_1 \ldots c_{l-1} \ldots \\ a'_0\ a'_1 \ldots a'_{n-1}\ b'_0\ b'_1 \ldots b'_{m-1}\ c'_0\ c'_1 \ldots c'_{l-1} \ldots \end{pmatrix},$$

so ergiebt sich (vgl. auch § 36 Beispiel 8):

$$\bar{Q}PQ = (a'_0, a'_1 \ldots a'_{n-1})\,(b'_0, b'_1 \ldots b'_{m-1})\,(c'_0, c'_1 \ldots c'_{l-1}) \ldots,$$

oder man erhält $\bar{Q}PQ$ dadurch, dafs man in den Cyklen von P die Permutation Q ausführt.

---

§ 39. Gruppen von Permutationen dreier Elemente.

Schreibt man die sämtlichen $3! = 6$ Permutationen von drei Elementen 0, 1, 2 hin, so findet man aufser der identischen 3 Transpositionen

$$(1, 2), \quad (0, 2), \quad (0, 1)$$

und 2 Cyklen dritter Ordnung

$$(0, 1, 2), \quad (0, 2, 1).$$

Wir wollen nun die sämtlichen Gruppen aufstellen, die sich aus ihnen bilden lassen, und behaupten, dafs in der vollständigen Gruppe nur 4 verschiedene Gruppen enthalten sind, nämlich eine Gruppe G dritter und drei Gruppen H_0, H_1, H_2 zweiter Ordnung. Die Gruppe G wird gebildet aufser der identischen Permutation von den zwei Cyklen dritter Ordnung, ist also darstellbar durch

$$G = \{1, (0, 1, 2), (0, 2, 1)\},$$

während die Gruppen H_0, H_1, H_2 aufser der identischen Permutation noch je eine Transposition enthalten, so dafs

$$H_0 = \{1, (1, 2)\}, \quad H_1 = \{1, (0, 2)\}, \quad H_2 = \{1, (0, 1)\}$$

ist.

Dafs die hingeschriebenen Permutationen wirklich Gruppen konstituieren, läfst sich leicht zeigen. Es mufs nur noch bewiesen werden, dafs es keine andern Gruppen giebt, die in der vollständigen enthalten sind. Da eine solche Gruppe nur dreigliedrige Cyklen und Transpositionen enthalten kann, so haben wir folgende Möglichkeiten zu beachten:

1. Enthält die Gruppe einen dreigliedrigen Cyklus, so enthält sie auch den andern, weil dieser durch Zusammensetzung aus ihm hervorgeht. Nämlich es ist

$$(a, b, c)^2 = (a, c, b).$$

Enthält nun die Gruppe keine andern Permutationen, so ist sie die Gruppe G. Sonst aber mufs sie mindestens noch eine Transposition enthalten. Dann aber enthält sie auch die beiden andern und ist also die vollständige Gruppe; denn es ist

$$(a, b, c)\ (a, b) = (b, c)$$
$$(a, b)\ (a, b, c) = (a, c).$$

2. Enthält die Gruppe keinen dreigliedrigen Cyklus, so kommen nur Transpositionen in Betracht. Ist nur eine Transposition vorhanden, so ist sie mit einer der Gruppen H_0, H_1, H_2 identisch. Enthält sie dagegen zwei verschiedene Transpositionen (a, b), (a, c), so enthält sie auch

$$(a, b)\ (a, c) = (a, b, c),$$

also einen dreigliedrigen Cyklus und ist daher, wie wir in 1. gesehen haben, die vollständige Gruppe.

Damit ist dargethan, dafs unsere Aufzählung der Gruppen vollständig war. Wir wollen nun noch auf eine einfache Darstellungsart der Gruppe eingehen und setzen

$$(1) \qquad S = (0, 1, 2), \qquad T = (1, 2).$$

Dann ist

$$(2) \qquad S^3 = 1, \qquad T^2 = 1$$

und

$$(3) \qquad ST = TS^2, \qquad S^2T = TS.$$

Mit Hülfe dieser Relationen ist es möglich, bei der Komposition von beliebig vielen Permutationen S und T immer eine solche Anordnung zu treffen, dafs auf eine Potenz von S eine Potenz von T folgt, so dafs $S^\sigma T^\tau$ die allgemeinste Form darstellt, auf die man jede Resultante bringen kann, wobei σ die Werte 0, 1, 2, τ die Werte 0, 1 annehmen kann. Die 6 Ausdrücke, die man so bilden kann, sind aber auch alle von einander verschieden, denn wäre $S^\sigma T^\tau = S^{\sigma'} T^{\tau'}$, so müfste $S^{\sigma - \sigma'} = T^{\tau - \tau'}$ sein; das ist aber nur möglich, wenn $\sigma \equiv \sigma' \bmod 3$, $\tau \equiv \tau' \bmod 2$ ist.

Es ergiebt sich hieraus, dafs die 6 verschiedenen Ausdrücke die sämtlichen Permutationen darstellen, was sich natürlich noch durch direkte Ausrechnung zeigt, wobei folgende Tabelle entsteht:

$$S = (0, 1, 2), \quad S^2 = (0, 2, 1),$$
$$T = (1, 2), \quad ST = (0, 2), \quad S^2T = (0, 1).$$

Die Gruppen stellen sich demnach in folgender Weise dar:

$$G = \{1, S, S^2\},$$
$$H_0 = \{1, T\}, \quad H_1 = \{1, ST\}, \quad H_2 = \{1, S^2T\},$$

und es mag der Übung des Lesers überlassen bleiben, die Gruppeneigenschaft der in geschweifte Klammern eingeschlossenen Permutationen auch mit Hülfe der oben dargelegten Relationen (3) zu bestätigen. Aus diesen beiden Relationen ergiebt sich, dafs G eine ausgezeichnete Untergruppe in der vollständigen Gruppe ist, während H_0, H_1, H_2 in dieser ein System conjugierter Gruppen bilden. Man kann sie nämlich auch in der Form darstellen

$$H_0 = \{1, T\}, \quad H_1 = \{1, S^{-1}TS\}, \quad H_2 = \{1, S^{-2}TS^2\},$$

woraus die Behauptung unmittelbar folgt.

---

§ 40. Die Alterngruppe.

Die Aufgabe, sämtliche aus den n! Permutationen von n Elementen zu bildenden Gruppen allgemein zu bestimmen, ist bis jetzt noch nicht gelöst. Nur für die einfachsten Werte von n hat man sämtliche Gruppen aufgestellt, doch existieren auch für ein beliebiges n eine Reihe allgemeiner Resultate. Wir wollen uns hier auf die sogenannte Alterngruppe beschränken und stellen dazu einige Sätze voran.

I) Jede Permutation von n Elementen kann aus Transpositionen zusammengesetzt werden und zwar speziell aus (n — 1) Transpositionen mit einem beliebigen gemeinsamen Element.

Da jede Permutation in Cyklen zerlegbar ist (§ 2), so braucht die Behauptung nur für solche bewiesen zu werden. Nun stellt aber (§ 36)

$$(a_1, a_2, .. a_m) = (a_1, a_2)(a_1, a_3) \dots (a_1, a_{m-1})$$

die Zerlegung eines beliebigen Cyklus in Transpositionen dar. Weil nun aber

$$(a_i, a_k) = (a_0, a_i)(a_0, a_k)(a_0, a_i)$$

ist, so kann man jede Transposition, in der ein beliebiges Element a_0 nicht vorkommt, durch solche ersetzen, in denen a_0 das eine Element ist. Sind $a_1, .. a_{n-1}$ die übrigen Elemente, so lassen sich also sämtliche Permutationen aus den $(n-1)$ Transpositionen

$$(a_0, a_1), (a_0, a_2) \dots (a_0, a_{n-1})$$

zusammensetzen.

Enthält eine Gruppe alle diese Transpositionen, so enthält sie alle überhaupt möglichen Permutationen von n Elementen und ist daher die vollständige Gruppe.

Die Zerlegung in Transpositionen ist jedoch keine eindeutige, wovon man leicht sich durch Beispiele überzeugen kann.

Für eine spätere Anwendung (§ 47) ist es wichtig zu zeigen, dafs man statt der oben genannten $(n-1)$ Transpositionen mit demselben gemeinschaftlichen Element, auch $(n-1)$ Transpositionen

$$(a_0, a_1), (a_1, a_2), \dots (a_{n-2}, a_{n-1}),$$

von denen je zwei aufeinanderfolgende ein Element gemeinsam haben, benutzen kann. Es ist allgemein

$$(a_0, a_{i+1}) = (a_i, a_{i+1})(a_0, a_i)(a_i, a_{i+1}). \quad (i = 1, 2, .. n-2.)$$

Ist also (a_0, a_i) durch die obengenannten Transpositionen darstellbar, so ist es auch (a_0, a_{i+1}). Da nun (a_0, a_1) selbst eine solche Transposition ist, so ist die Behauptung bewiesen.

II) Läfst sich eine aus m Cyklen ohne gemeinschaftliche Elemente bestehende und n Elemente enthaltende Permutation auch durch t Transpositionen darstellen, so ist $m + n + t$ stets eine gerade Zahl.

Dieser Satz trifft für 1 Transposition zu, da diese 1 Cyklus mit 2 Buchstaben darstellt. Auch für 2 Transpositionen ergiebt sich die Richtigkeit des Satzes. Denn

kommt noch eine Transposition hinzu, so kann diese mit der schon vorhandenen entweder beide Elemente oder eins oder keins gemein haben. Im ersten Falle entsteht die identische Permutation, bei der die Anzahl der zu vertauschenden Elemente und Cyklen gleich Null ist. Im zweiten Falle entsteht, wie

$$(a, b)(a, c) = (a, b, c)$$

lehrt, ein dreigliedriger Cyklus und ist also $m = 1$, $n = 3$, $t = 2$. Im dritten Falle endlich bilden die beiden Transpositionen zwei Cyklen, es ist $m = t = 2$, $n = 4$.

Wir wollen nun annehmen, der Satz sei für t Transpositionen bewiesen, und dann zeigen, dafs er auch noch gilt, wenn noch eine Transposition hinzutritt. Die Anzahl der Cyklen und der in ihnen vorkommenden Elemente der entstehenden Permutationen bezeichnen wir durch m' resp. n'. Wir unterscheiden wieder drei Hauptfälle:

1) Die Elemente der neu hinzukommenden Transposition kommen in keinem Cyklus vor, also ist $n' = n + 2$; dann wird $m' = m + 1$.

2) Nur ein Element der neuen Transposition kommt in den Cyklen vor, dann ist $n' = n + 1$. Wie dann

$$(a, a_1, .. a_{r-1})(a, b) = (a, a_1 \dots a_{r-1}, b)$$

zeigt, wird $m' = m$.

3) Beide Elemente der Transposition kommen in den Cyklen der Permutation vor. Wir unterscheiden, ob die beiden Elemente in demselben oder in verschiedenen Cyklen vorkommen. Tritt das Erstere ein, so lehren die Beziehungen:

$$(a, b)^2 = 1, \quad (a, a_1, .. a_{r-1}, b)(a, b) = (a, a_1, .. a_{r-1}),$$

$$(a, a_1 .. a_{r-1}, b, b_1 .. b_{s-1})(a, b) = (a, a_1 .. a_{r-1})(b, b_1 .. b_{s-1}),$$

dafs entweder $m' = m - 1$, $n' = n - 2$, oder $m' = m$, $n' = n - 1$, oder $m' = m + 1$, $n' = n$ wird. Trifft das Letztere zu, so ergiebt sich aus

$$(a, a_1 .. a_{r-1})(b, b_1 .. b_{s-1})(a, b) = (a, a_1 .. a_{r-1})(b, b_1 .. b_{s-1})$$

$$m' = m - 1, \; n' = n.$$

In allen diesen Fällen ist $t + 1 + m' + n'$ wieder eine gerade Zahl, wenn $t + m + n$ eine solche war. Damit ergiebt sich die allgemeine Richtigkeit des Satzes, den wir zu beweisen hatten, und dem wir auch die Form

$$t \equiv m + n \bmod 2$$

geben können. Da die Anzahlen m und n bei jeder Permutation eindeutig bestimmt sind, so ist auch t bis auf Vielfache von 2 bestimmt. Also folgt der Satz:

III) Alle Permutationen zerfallen in zwei Klassen, je nachdem sie durch eine gerade oder ungerade Anzahl von Transpositionen dargestellt werden können.

Und unmittelbar folgt hieraus:

IV) Alle durch eine gerade Anzahl von Transpositionen darstellbaren Permutationen bilden eine Gruppe.

Denn: wenn zwei Permutationen A und B durch eine gerade Anzahl von Transpositionen dargestellt werden, so gilt dasselbe von A B, und A B gehört also wieder der Gruppe an.

Diese definierte Gruppe nennt man die Alterngruppe. Um sie weiter zu untersuchen, beweisen wir zunächst den folgenden Satz:

V) Jede Permutation von n Elementen läfst sich darstellen durch (n — 2) Cyklen dritter Ordnung mit zwei gemeinsamen Elementen und eine Transposition dieser beiden.

Bezeichnen wir die n Elemente durch $a, b, c_1, c_2, \ldots c_{n-2}$, so können die Cyklen und die Transposition durch

$$S_i = (a, b, c_i) \qquad (i = 1, 2, \ldots n-2)$$
$$T = (a, b)$$

dargestellt werden. Da jede Permutation aus der Transposition T und den folgenden

$$(a, c_1), \ldots . (a, c_{n-2})$$

zusammengesetzt werden kann, so ist nur noch zu zeigen, dafs diese letzteren in der angegebenen Weise zu zerlegen sind. Das ergiebt sich aber aus

$$(a, c_i) = (a, b)(a, b, c_i) = T S_i. \qquad (i = 1, 2 \ldots n-2).$$

Es läfst sich ferner zeigen, dafs man die Darstellung jeder Permutation P so einrichten kann, dafs T nur ein einziges Mal, und zwar zuletzt gebraucht wird. Nämlich mit Hilfe der Gleichungen

$$S_1^3 = 1, \quad TS_1 = S_1^2 T,$$

aus denen

$$T S_1^2 = S_1 T$$

folgt, kann man T an das Ende bringen. Da $T^2 = 1$ ist, so wird T fehlen oder vorhanden sein, je nachdem es in einer geraden oder ungeraden Anzahl vorkommt.

Da jeder dreigliedrige Cyklus in zwei Transpositionen zerlegt werden kann, so ist jede Permutation, in der T fehlt, durch eine gerade, aber jede, bei deren Darstellung T vorhanden ist, durch eine ungerade Anzahl von Transpositionen darstellbar. Daraus folgt dann:

VI) Alle Permutationen der Alterngruppe lassen sich durch Zusammensetzung von $(n - 2)$ Cyklen dritter Ordnung mit zwei gemeinsamen Elementen darstellen, und umgekehrt gehören alle so darstellbaren Permutationen der Alterngruppe an. Die Ordnung der Alterngruppe ist $\frac{n!}{2}$.

Das Letztere folgt daraus, dafs man alle Permutationen der vollständigen Gruppe erhält, wenn man die Permutationen der Alterngruppe mit T zusammensetzt. Hierbei ist es gleichgültig, ob man T überall als rechte oder überall als linke Komponente gebraucht.

Ist P eine Permutation der Alterngruppe, so ist auch $\bar{T} P T = T P T$ eine solche.

VII) Daher ist die Alterngruppe eine ausgezeichnete Untergruppe der vollständigen Gruppe.

Dieser Satz läfst sich umkehren in der Form:

VIII) Jede ausgezeichnete Untergruppe der vollständigen Gruppe, die einen dreigliedrigen Cyklus als Element enthält, ist die Alterngruppe.

Denn wenn die ausgezeichnete Untergruppe den dreigliedrigen Cyklus

$$S = (a, b, c)$$

enthält und P eine beliebige Permutation ist, so mufs sie auch $\overline{P}SP$ enthalten. Wählt man aber

$$P = \begin{pmatrix} a & b & c & \dots \\ a' & b' & c' & \dots \end{pmatrix},$$

so wird $\overline{P}SP = (a', b', c')$ und kann also, da a', b', c' ganz beliebig sind, jeden dreigliedrigen Cyklus darstellen, so dafs die sämtlichen Permutationen der Alterngruppe in der ausgezeichneten Gruppe enthalten sind.

Gehört nun P nicht der Alterngruppe an, so ist

$$P' = \begin{pmatrix} a & b & c & \dots \\ a' & b' & c' & \dots \end{pmatrix} (a', b') = \begin{pmatrix} a & b & c & \dots \\ b' & a' & c' & \dots \end{pmatrix}$$

eine Permutation der Alterngruppe, und es ist daher

$$\overline{P'}SP' = (a', c', b') = (a', b', c')^2$$

ein beliebiger Cyklus von drei Elementen. Man kann also dieselben Schlüsse statt auf die vollständige Gruppe auch auf die Alterngruppe selbst anwenden und erhält dadurch den Satz:

IX) Es giebt keine ausgezeichnete Untergruppe in der Alterngruppe, die einen Cyklus von drei Elementen enthielte.

Denn eine solche wäre mit der Alterngruppe identisch, also keine eigentliche Untergruppe derselben.

---

§ 41. Zerlegung der Alterngruppe von vier Elementen.

Die sämtlichen 24 Permutationen von vier Elementen 0, 1, 2, 3 zerfallen aufser der identischen in folgende Arten:

6 Transpositionen:

(0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 3),

8 dreigliedrige Cyklen:

(1, 2, 3), (0, 2, 3), (0, 1, 3), (0, 1, 2)
(1, 3, 2), (0, 3, 2), (0, 3, 1), (0, 2, 1),

6 viergliedrige Cyklen:

$$(0, 1, 2, 3),\ (0, 1, 3, 2),\ (0, 2, 3, 1),$$
$$(0, 2, 1, 3),\ (0, 3, 1, 2),\ (0, 3, 2, 1),$$

und endlich drei Transpositionspaare:

$$(0, 1)\ (2, 3),\ (0, 2)\ (1, 3),\ (0, 3)\ (1, 2).$$

Wir wollen nun nicht, wie in § 39 die allgemeine Aufgabe lösen, alle Untergruppen zu bestimmen, die in der vollständigen Gruppe enthalten sind, sondern uns darauf beschränken, aus dieser eine Reihe von Gruppen auszusondern, von denen jede folgende in der vorhergehenden als ausgezeichnete Untergruppe enthalten ist, weil diese allein bei der Auflösung der Gleichung vierten Grades später in Betracht kommen. Aus der Untersuchung des vorhergehenden Paragraphen wissen wir bereits, dafs die Alterngruppe G die Gruppe von höchstem Grade ist, die in der vollständigen Gruppe als ausgezeichnete enthalten ist. Wir brauchen also nur diese selbst weiter zu zerlegen. Vorab bemerken wir aber, dafs sie nicht die 8 Transpositionen, ebenso auch nicht die 6 viergliedrigen Cyklen enthält, die, wie aus

$$(a, b, c, d) = (a, b)\ (a, c)\ (a, d)$$

hervorgeht, sich in drei Transpositionen zerlegen lassen; dafs dagegen alle andern Permutationen, also aufser der identischen die 8 dreigliedrigen Cyklen und 3 Transpositionspaare der Alterngruppe G zugehören.

Eine in ihr enthaltene ausgezeichnete Untergruppe kann nun keinen dreigliedrigen Cyklus besitzen, wie wir im vorigen Paragraphen (Satz IX) bewiesen haben. Somit bleibt nur noch übrig anzunehmen, dafs sie ein Transpositionspaar

$$S = (a, b)\ (c, d)$$

enthält. Dann enthält sie, da

$$T = (a, b, c)$$

eine Permutation in G ist, auch das andere

$$T^{-1}\,S\,T = (a, d)\ (b, c) = S',$$

und weil

$$S\,S' = (a, c)\ (b, d) = S''$$

ist, das dritte Transpositionspaar, die in Verbindung mit der identischen Permutation eine Gruppe und auch eine ausgezeichnete bilden, weil jedes Transpositionspaar durch einen dreigliedrigen Cyklus in ein anderes, durch ein Transpositionspaar in sich selbst transformiert wird.

Wir setzen nun

$$S_1 = (0, 1)(2, 3), \quad S_2 = (0, 2)(1, 3),$$

so dafs

$$S_1^2 = 1, \ S_2^2 = 1, \ S_1 S_2 = S_2 S_1 = (0, 3)(1, 2)$$

ist und die genannte Gruppe durch

$$H = \{1, S_1, S_2, S_1 S_2\}$$

dargestellt werden kann. Sie besitzt, wie leicht zu erkennen ist, drei Untergruppen, die alle ausgezeichnet in H sind, nämlich die Gruppen

$$K_1 = \{1, S_1\}, K_2 = \{1, S_2\}, K_3 = \{1, S_1 S_2\}.$$

Führt man nun noch

$$T = (1, 2, 3)$$

ein, so kann man mit Hülfe der Relationen

$$T^3 = 1, \ S_1 T = T S_2, \ S_2 T = T S_1 S_2$$

die Permutationen der Alterngruppe in der Form

$$S_1^{\sigma_1} S_2^{\sigma_2} T^{\tau}$$

darstellen, wo σ_1, σ_2 die Werte 0, 1, τ die Werte 0, 1, 2 annehmen kann, was der Leser sich selbst ableiten mag, ebenso, dafs wenn nun noch eine nicht in G enthaltene Permutation, etwa

$$U = (1, 2)$$

zu Hülfe genommen wird, sich auf Grund der weiteren Relationen

$$U^2 = 1, \quad S_1 U = U S_2, \quad S_2 U = U S_1, \quad T U = U T^2$$

eine Darstellung der vollständigen Gruppe in der Form

$$S_1^{\sigma_1} S_2^{\sigma_2} T^{\tau} U^{\omega}$$

ergiebt, wo ω dann noch die Werte 0, 1 annehmen kann.

Es läſst sich hieraus auch leicht erkennen, daſs H nicht nur eine ausgezeichnete Untergruppe der Alterngruppe G, sondern auch der vollständigen Gruppe ist, denn es ist

$$S_1 = U^{-1} S_2 U, \quad S_2 = U^{-1} S_1 U.$$

Die drei Gruppen K_1, K_2, K_3, die in H als ausgezeichnete Untergruppen enthalten sind, stellen dagegen in Bezug auf die Alterngruppe und die vollständige Gruppe ein vollständiges System von drei konjugierten Untergruppen dar, wie man ebenfalls leicht zeigen kann.

---

§ 42. Einfachheit der Alterngruppe.

Während die Alterngruppe für drei Elemente schon aus dem Grunde, weil ihr Grad eine Primzahl ist, eine einfache Gruppe sein muſs, hat sie sich für vier Elemente als zusammengesetzt erwiesen. Es ist nun von groſser Wichtigkeit, daſs sie mit alleiniger Ausnahme dieses Falles eine einfache Gruppe darstellt, obgleich ihre Gradzahl für mehr als vier Elemente zusammengesetzt ist; denn diese Thatsache bildet die gruppentheoretische Grundlage für die algebraische Unauflösbarkeit der allgemeinen Gleichungen von höherem als dem vierten Grade.

Bezeichnen wir mit S eine Permutation, die in einer ausgezeichneten Untergruppe der Alterngruppe enthalten sei, mit T eine der Alterngruppe selbst, so gehört der ersteren auch $T^{-1}ST$ und daher auch $S^{-1}T^{-1}ST$ an. Ist

$$S = \begin{pmatrix} 1 & 2 & \ldots & n \\ i_1 & i_2 & \ldots & i_n \end{pmatrix}, \quad T = (a, b, c),$$

so ist $S^{-1}T^{-1}S = (i_c, i_b, i_a)$. Es läſst sich aber $T = (a, b, c)$ stets so wählen, daſs $S^{-1}T^{-1}S$ und T entweder kein einziges oder ein Element oder zwei in entgegengesetzter Reihenfolge gemein haben, und daſs in den beiden letzten Fällen nach den Formeln

$$(a, b, c)\,(a, d, e) = (a, b, c, d, e)$$
$$(a, b, c)\,(b, a, d) = (b, c, d)$$

$S^{-1}T^{-1}ST$ ein Cyklus von 5 oder von 3 Elementen wird. Wir unterscheiden folgende 4 Fälle:

1) S enthalte ein Transpositionspaar (a, a′) (b, b′). Da ein fünftes Element c existiert, so wählen wir, wenn dieses durch S nicht geändert wird, $T = (a, b, c)$ und erhalten $S^{-1}T^{-1}S = (c, b', a')$, $S^{-1}T^{-1}ST = (a, b, c, b', a')$.

2) Giebt es dagegen kein Element, das durch S unverändert bleibt, und enthält S die Transpositionen (a, a′), (b, b′), (c, c′), so ergiebt sich, wenn $T = (a, b, c)$ angenommen wird, $S^{-1}T^{-1}S = (c', b', a')$, $S^{-1}T^{-1}ST = (c', b', a')(a, b, c)$, was unter die folgende Annahme fällt.

3) S enthalte aufser einem dreigliedrigen Cyklus (a, a′, a″) noch einen andern solchen (b, b′, b″) oder eine Transposition (b, b′). Bei $T = (a, a', b)$ wird dann $S^{-1}T^{-1}S = (a', b', a'')$, $S^{-1}T^{-1}ST = (a, a', b', a'', b)$.

4) S enthalte einen Cyklus von mehr als drei Elementen (a, a′, a″, a‴ . . .). Ist dann $T = (a, a', a'')$, so folgt $S^{-1}T^{-1}S = (a''', a'', a')$ und $S^{-1}T^{-1}ST = (a, a', a''')$.

Der Fall 4) ist aber auf die weitere Untersuchung der Fälle 1), 3) anwendbar und andere Fälle sind, abgesehen davon, dafs S ein dreigliedriger Cyklus sein kann, unmöglich, da S keine Transposition sein darf, weil die Alterngruppe eine solche nicht enthält. Es ergiebt sich also, dafs die Gruppe immer einen dreigliedrigen Cyklus enthält. Daraus folgt (§ 40, IX), dafs sie nur mit der Alterngruppe identisch sein kann, die also keine eigentliche ausgezeichnete Untergruppe enthält und daher eine einfache Gruppe darstellt.

VI. Abschnitt.

Determinanten.

---

§ 43. Systeme zweier linearer Gleichungen mit zwei Unbekannten.

Wir beschäftigen uns in diesem Abschnitt allgemein mit der Auflösung eines Systems linearer Gleichungen mit mehreren Unbekannten, wollen jedoch dieses Problem nicht sofort in seiner Allgemeinheit in Angriff nehmen, sondern erst an speziellen Fällen die allgemeine Auffassung entwickeln.

Bei der Auflösung des Gleichungssystems

$$(1) \qquad \begin{cases} a_1 x + b_1 y + c_1 = 0 \\ a_2 x + b_2 y + c_2 = 0 \end{cases}$$

kann man bekanntlich folgendes Verfahren einschlagen:

I. Man multipliziert die Seiten der Gleichungen mit b_2, b_1 darauf mit a_2, a_1 und erhält dann durch Subtraktion die beiden Gleichungen

$$(2) \qquad \begin{cases} (a_1 b_2 - a_2 b_1) x + c_1 b_2 - c_2 b_1 = 0 \\ (a_1 b_2 - a_2 b_1) y + a_1 c_2 - a_2 c_1 = 0, \end{cases}$$

von denen die erste nur x, die zweite nur y als Unbekannte enthält, und die sofort, wenn

$$a_1 b_2 - a_2 b_1 \neq 0$$

ist, die Lösungen

$$(3) \qquad \begin{cases} x = \dfrac{b_1 c_2 - b_2 c_1}{a_1 b_2 - a_2 b_1} \\ y = \dfrac{a_2 c_1 - a_1 c_2}{a_1 b_2 - a_2 b_1} \end{cases}$$

ergeben. Dafs diese eindeutig bestimmten Werte nun aber auch wirklich Lösungen sind, kann man entweder durch Einsetzen in die Gleichungen (1) bestätigen, oder aber besser dadurch erkennen, dafs man vom Gleichungssystem (3) auf (2) und schliefslich auf (1) rückwärts schliefst. Der erste Schlufs ist unmittelbar zu machen; multipliziert man darauf die Gleichungen (2) mit a_1, b_1 resp. a_2, b_2, und addiert, so erhält man

$$(4)\qquad \begin{cases}(a_1 b_2 - a_2 b_1)\,(a_1 x + b_1 y + c_1) = 0\\ (a_1 b_2 - a_2 b_1)\,(a_2 x + b_2 x + c_2) = 0,\end{cases}$$

woraus sich infolge der Voraussetzung $a_1 b_2 - a_2 b_1 \neq 0$ sofort das System (1) ergiebt.

II. Unsere Schlüsse waren aber durchaus an die Voraussetzung $a_1 b_2 - a_2 b_1 \neq 0$ gebunden. Lassen wir diese nun fallen, so geht aus (2) sofort hervor, dafs dann zugleich

$$a_1 b_2 - a_2 b_1 = 0, \quad b_1 c_2 - b_2 c_1 = 0, \quad a_1 c_2 - a_2 c_1 = 0$$

sein mufs. Die Gleichungen (2) werden nun durch jeden beliebigen Wert von x und y befriedigt; da wir aber von diesen auf (1) nur auf Grund der jetzt nicht mehr zutreffenden Annahmen geschlossen haben, so dürfen wir nicht behaupten, dafs den Gleichungen (1) durch jeden beliebigen Wert von x und y Genüge geleistet wird. Das ist sicher nicht der Fall, sobald nur eine der vier Gröfsen

$$\begin{matrix} a_1, b_1 \\ a_2, b_2 \end{matrix}$$

von Null verschieden ist. Ist z. B. $a_1 \neq 0$, so folgt aus der ersten Gleichung in (1)

$$x = -\frac{b_1}{a_1} y - \frac{c_1}{a_1},$$

dafs x vollständig bestimmt ist, sobald man für y einen beliebigen Wert angenommen hat. Dafs man dies nun aber wirklich thun darf, geht daraus hervor, dafs die zweite Gleichung in (1) befriedigt wird, wenn man den Wert für x einsetzt, denn die linke Seite geht dann über in $\frac{(a_2 b_1 - b_2 a_1)\,y + c_2 a_1 - a_2 c_1}{a_1}$, was unseren Annahmen zufolge verschwindet und zwar für jeden beliebigen Wert

von y. Man kann daher y beliebig wählen, worauf x völlig bestimmt ist.

Eben dieselben Schlüsse gelten auch dann, wenn $a_1 = 0$ ist, dagegen eine der noch übrigen Gröfsen nicht verschwindet, die dann an Stelle von a_1 in unseren Schlüssen zu treten hat. Es ergiebt sich somit, dafs unter unserer Voraussetzung immer unendlich viele Wertepaare die beiden vorgelegten Gleichungen (1) befriedigen, dafs man einer der beiden Unbekannten — entweder nur x oder nur y oder auch beiden — willkürliche Werte erteilen kann, worauf der Wert der andern völlig bestimmt ist.

III. Um nun noch unsere Betrachtung zum Abschlufs zu bringen, fügen wir hinzu, dafs, wenn alle vier Gröfsen

$$\begin{matrix} a_1, b_1 \\ a_2, b_2 \end{matrix}$$

verschwinden, auch c_1 und c_2 verschwinden müssen, und dafs das Gleichungssystem (1) dann durch jeden beliebigen Wert von x und von y befriedigt wird.

Blicken wir nun noch einmal auf diese Entwickelung zurück, so zeigt es sich von Wichtigkeit, dafs man nicht nur vom Gleichungssystem auf die Lösungen zu schliefsen hat, sondern dafs sich aus den Lösungen das ursprüngliche System wieder aufbauen lassen mufs. Die Lösungen werden nun auch von zwei Gleichungen wie in (2) oder (3), oder von einer Gleichung wie bei II gegeben, während bei III keine Gleichung die Werte von x und y einschränkt. Um nun den Zusammenhang zwischen dem ursprünglichen Gleichungssystem und dem Lösungssystem zu untersuchen, ist es zweckmäfsig, für die auf den linken Seiten der Gleichungen auftretenden linearen Funktionen Bezeichnungen einzuführen. Wir setzen

$$(5) \qquad \begin{aligned} f_1 &= a_1 x + b_1 y + c_1 \\ f_2 &= a_2 x + b_2 y + c_2 \end{aligned}$$

$$(6) \qquad \begin{aligned} \varphi_1 &= (a_1 b_2 - a_2 b_1) x + (c_1 b_2 - c_2 b_1) \\ \varphi_2 &= (a_1 b_2 - a_2 b_1) y + (a_1 c_2 - a_2 c_1), \end{aligned}$$

so dafs das ursprüngliche Gleichungssystem (1) durch $f_1 = 0$, $f_2 = 0$ dargestellt, das Lösungssystem unter der Voraussetzung $a_1 b_2 - a_2 b_1 \neq 0$ aus $\varphi_1 = 0$, $\varphi_2 = 0$ abgeleitet

werden kann. Man erkennt nun aus der Art der Ableitung der Gleichungen (2) aus (1) zunächst, dafs

$$(7) \qquad \begin{aligned} \varphi_1 &= b_2 f_1 - b_1 f_2 \\ \varphi_2 &= - a_2 f_1 + a_1 f_2 \end{aligned}$$

ist, dafs sich also φ_1 und φ_2 linear und homogen durch f_1, f_2 ausdrücken lassen und aus diesem Grunde verschwinden müssen, sobald $f_1 = 0$, $f_2 = 0$ ist. Aber auch das Umgekehrte findet statt, wenn $a_1 b_2 - a_2 b_1 \neq 0$ ist, und in der That lassen sich auch umgekehrt f_1 und f_2 linear und homogen durch f_1, f_2 ausdrücken, denn aus (7) ergiebt sich

$$(8) \qquad \begin{aligned} f_1 &= \frac{a_1}{a_1 b_2 - a_2 b_1} \varphi_1 + \frac{b_1}{a_1 b_2 - a_2 b_1} \varphi_2 \\ f_2 &= \frac{a_2}{a_1 b_2 - a_2 b_6} \varphi_1 + \frac{b_2}{a_1 b_2 - a_2 b_1} \varphi_2, \end{aligned}$$

wobei die Koeffizienten von endlichen Werten sind. Wir haben nun früher (§ 17) Modulsysteme aus ganzen Zahlen betrachtet, und als hinreichende und notwendige Bedingung für die Äquivalenz zweier solcher Modulsysteme kennen gelernt, dafs sich jedes Element des einen linear und homogen durch die des andern ausdrücken lassen mufs (§ 18). Diesen Umstand wollen wir jetzt zu einer Verallgemeinerung des Begriffes der Modulsysteme benutzen und von zwei Systemen von Gröfsen, von denen die des einen sich linear und homogen durch die des andern Systems und umgekehrt ausdrücken lassen, sagen, dafs sie äquivalente Modulsysteme darstellen. Dann können wir den Zusammenhang der Funktionen $f_1, f_2, \varphi_1, \varphi_2$ nicht kürzer ausdrücken als durch die Äquivalenz

$$(f_1, f_2) = (\varphi_1, \varphi_2).$$

Dies gilt jedoch nur, wenn $a_1 b_2 - a_2 b_1$ nicht verschwindet. Ist $a_1 b_2 - a_2 b_1 = 0$, so sind die Gleichungen (7) zwar auch noch richtig, aber die Gleichungen (8) widersinnig. Nun haben wir aber unter II gesehen, dafs unter der Voraussetzung $a_1 \neq 0$ die Gleichung $f_1 = 0$ als Lösung betrachtet werden kann. An Stelle der Gleichungen (8) kann man in diesem Falle eine einzige setzen. Aus der Gleichung

$$f_2 = \frac{a_2}{a_1} f_1 + \frac{a_1 b_2 - a_2 b_1}{a_1} y + \frac{c_1 a_2 - a_1 c_2}{a_1}$$

folgt, weil $a_1 b_2 - a_2 b_1 = 0$ ist,

$$(9) \qquad f_2 = \frac{a_2}{a_1} f_1 + \frac{c_1 a_2 - a_1 c_2}{a_1},$$

und nun ergiebt sich, dafs

$$(f_1, f_2) = (f_1, c_1 a_2 - a_1 c_2)$$

ist, da f_2 einerseits linear und homogen durch f_1 und $c_1 a_2 - a_2 c_1$, andererseits $c_1 a_2 - a_1 c_2$ linear und homogen durch f_1 und f_2 mit endlichen Koeffizienten ausgedrückt werden kann. Wenn nun die Elemente f_1, f_2 des ersten Modulsystems verschwinden, so verschwinden auch die des zweiten und umgekehrt. Die hinreichenden und notwendigen Bedingungen für das Vorhandensein einer Lösung sind also

$$f_1 = 0, \ a_1 c_2 - a_2 c_1 = 0,$$

und die früher noch abgeleitete Bedingung $b_1 c_2 - b_2 c_1 = 0$ ist auch hiervon eine Folge, da

$$a_1 (b_1 c_2 - b_2 c_1) = - c_1 (a_1 b_2 - a_2 b_1) + b_1 (a_1 c_2 - a_2 c_1)$$

ist.

Sind endlich alle Gröfsen

$$\begin{matrix} a_1, & b_1 \\ a_2, & b_2 \end{matrix}$$

gleich Null, so ergiebt sich leicht

$$(f_1, f_2) = (c_1, c_2)$$

und daraus als notwendige und hinreichende Bedingung für die Existenz von Lösungen, dafs auch c_1 und c_2 verschwinden müssen.

Haben wir so den inneren Zusammenhang zwischen dem ursprünglichen Gleichungssystem und den Lösungen mit Hilfe einer Erweiterung des Begriffes der Modulsysteme dargestellt, so wollen wir zum Schlusse die Aufmerksamkeit darauf lenken, dafs das verschiedenartige Verhalten des Gleichungssystemes nicht sowohl von den Koeffizienten der Gleichungen, als von gewissen Verbindungen abhängig ist; es trat hierbei der Ausdruck $a_1 b_2 - a_2 b_1$ auf, dessen Verschwinden oder Nichtverschwinden den Fall I von den übrigen absondert. Einen solchen Ausdruck schreibt man abgekürzt

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

und nennt ihn eine **Determinante** zweiter Ordnung. Er tritt aber auch in den Lösungen auf, denen man bei nichtverschwindender Determinante die Form

$$x = -\frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

$$y = -\frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

geben kann, wie leicht zu sehen ist.

---

§ 44. Systeme von drei linearen Gleichungen mit drei Unbekannten.

Die Gesichtspunkte, zu denen wir bei der Behandlung des Gleichungssystems des vorigen Paragraphen zuletzt gelangt sind, wollen wir jetzt, wo das System $f_1 = 0$, $f_2 = 0$, $f_3 = 0$ dreier linearer Gleichungen mit drei Unbekannten untersucht werden soll, schon von vornherein zur Geltung kommen lassen. Es sei hierbei

$$\begin{aligned} f_1 &= a_1 x + b_1 y + c_1 z + d_1 \\ f_2 &= a_2 x + b_2 y + c_2 z + d_2 \\ f_3 &= a_3 x + b_3 y + c_3 z + d_3. \end{aligned}$$

I. Sind alle neun Gröfsen

$$\begin{matrix} a_1, & b_1, & c_1 \\ a_2, & b_2, & c_2 \\ a_3, & b_3, & c_3 \end{matrix}$$

gleich Null, so ist das Modulsystem (f_1, f_2, f_3) äquivalent dem System (d_1, d_2, d_3). Hieraus ergiebt sich, dafs die Gleichungen $f_1 = 0$, $f_2 = 0$, $f_3 = 0$ nur befriedigt werden können, wenn auch $d_1 = 0$, $d_2 = 0$, $d_3 = 0$ ist und zwar durch ganz beliebige und von einander unabhängige Werte für x, y, z.

Schliefsen wir nun das gleichzeitige Verschwinden der genannten Gröfsen aus, so können wir annehmen, dafs $a_1 \neq 0$ ist, ohne der Allgemeinheit Abbruch zu thun, wenn wir uns vorbehalten, die Bezeichnung der Funktionen und Variabeln nachträglich zu ändern. Wären z. B. die drei Gröfsen der ersten Reihe gleich Null, ebenso auch noch a_2, so würde man f_2 an Stelle von f_1 und y an Stelle von x setzen u. s. w.

Halten wir also an der Annahme $a_1 \neq 0$ fest und bilden nun die Funktionen

$$f_2' = f_2 - \frac{a_2}{a_1} f_1$$

$$f_3' = f_3 - \frac{a_3}{a_1} f_1,$$

so enthalten diese die Gröfse x nicht mehr und haben also die Form

$$f_2' = b_2' y + c_2' z + d_2'$$
$$f_3' = b_3' y + c_3' z + d_3',$$

wo

$$b_2' = b_2 - \frac{a_2}{a_1} b_1, \; c_2' = c_2 - \frac{a_2}{a_1} c_1, \; d_2' = d_2 - \frac{a_2}{a_1} d_1,$$

$$b_3' = b_3 - \frac{a_3}{a_1} b_1, \; c_3' = c_3 - \frac{a_3}{a_1} c_1, \; d_3' = d_3 - \frac{a_3}{a_1} d_1$$

gesetzt ist. Da f_2', f_3' linear aus f_1, f_2, f_3 gebildet sind, andererseits aber, wie die Gleichungen

$$f_2 = f_2' + \frac{a_2}{a_1} f_1$$

$$f_3 = f_3' + \frac{a_3}{a_1} f_1$$

zeigen, f_2 und f_3 sich linear durch f_1, f_2', f_3' ausdrücken lassen, so ist das Modulsystem (f_1, f_2, f_3) äquivalent dem System (f_1, f_2', f_3'). Um dieses letztere zu behandeln, erinnern wir daran, dafs das Teilsystem (f_2', f_3') von der Form ist, wie wir es schon im vorigen Paragraphen angetroffen haben. Wir haben demgemäfs noch drei verschiedene Fälle zu unterscheiden, die wir jetzt mit II, III, IV bezeichnen wollen.

II. Wir nehmen an, dafs die vier Gröfsen

$$\begin{matrix} b_2', & c_2' \\ b_3', & c_3' \end{matrix}$$

oder die Determinanten

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & b_1 \\ a_3 & b_3 \end{vmatrix}, \quad \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix}$$

verschwinden. Wie leicht abgeleitet werden kann, verschwinden dann alle Determinanten zweiter Ordnung, die man aus der Matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

bilden kann. Dann ist

$$(f_2', f_3') = (d_2', d_3'),$$

also

$$(f_1, f_2, f_3) = (f_1, d_2', d_3').$$

Hieraus ist nun sofort ersichtlich, dafs die Bedingung der Lösbarkeit lautet: $d_2' = 0$, $d_3' = 0$, dafs dann aber y und z jeden beliebigen Wert annehmen können, während x durch die Gleichung

$$x = -\frac{b_1}{a_1} y - \frac{c_1}{a_1} z - \frac{d_1}{a_1}$$

darauf völlig bestimmt ist. Beachten wir die Ausdrücke für d_2' und d_3', so müssen in diesem Falle überhaupt alle Determinanten zweiter Ordnung der Matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix}$$

verschwinden.

Wir nehmen nun ferner an, dafs eine der zuerst genannten Gröfsen nicht verschwindet, nämlich b_2'. Da wir noch über die Bezeichnung von f_2' und f_3', sowie y und z willkürlich verfügen können, so ist diese Annahme gestattet. Bilden wir nun

$$f_3'' = f_3' - \frac{b_3'}{b_2'} f_2',$$

so ist f_3'' nur noch von z abhängig und in der Form

$$f_3'' = c_3'' z + d_3''$$

darstellbar, wobei

$$c_3'' = c_3' - \frac{b_3'}{b_2'} c_2', \qquad d_3'' = d_3' - \frac{b_3'}{b_2'} d_2'$$

gesetzt ist. Diese Gröfsen lassen sich nun durch die ursprünglichen a_1, b_1, c_1, d_1 u. s. w. in folgender Form ausdrücken:

$$c_3'' = \frac{1}{b_2'} (a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1)$$

$$d_3'' = \frac{1}{b_2'} (a_1 b_2 d_3 - a_1 b_3 d_2 + a_2 b_3 d_1 - a_2 b_1 d_3 + a_3 b_1 d_2 - a_3 b_2 d_1).$$

Hier zeigt es sich nun sofort, dafs es wesentlich zur Abkürzung beiträgt, wenn wir für die in den Klammern eingeschlossenen Ausdrücke, die durchaus analog gebaut sind und Determinanten dritter Ordnung genannt werden, eine Bezeichnung wählen, nämlich setzen:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1,$$

dann wird

$$\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} = a_1 b_2 d_3 - a_1 b_3 d_2 + a_2 b_3 d_1 - a_2 b_1 d_3 + a_3 b_1 d_2 - a_3 b_2 d_1,$$

und wir haben

$$c_3'' = \frac{1}{b_2'} \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \qquad d_3'' = \frac{1}{b_2'} \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}.$$

Da f_3'' linear aus f_2' und f_3' gebildet worden ist, f_3' sich aber, wie aus der Gleichung

$$\mathfrak{f}_3' = \mathfrak{f}_3'' + \frac{\mathfrak{b}_3'}{\mathfrak{b}_2'}\,\mathfrak{f}_2'$$

hervorgeht, linear durch $\mathfrak{f}_2'$ und $\mathfrak{f}_3''$ darstellen läſst, so ist

$$(\mathfrak{f}_2', \mathfrak{f}_3') = (\mathfrak{f}_2', \mathfrak{f}_3''),$$

und somit folgt

$$(\mathfrak{f}_1, \mathfrak{f}_2, \mathfrak{f}_3) = (\mathfrak{f}_1, \mathfrak{f}_2', \mathfrak{f}_3'').$$

Das rechts stehende System können wir nun noch weiter vereinfachen. Da nämlich $\mathfrak{b}_2' \neq 0$ vorausgesetzt wurde, so kann man in

$$\mathfrak{f}_1' = \mathfrak{f}_1 - \frac{\mathfrak{b}_1}{\mathfrak{b}_2'}\,\mathfrak{f}_2'$$

eine Funktion $\mathfrak{f}_1'$ konstruieren, in der y nicht vorkommt, die also die Gestalt hat

$$\mathfrak{f}_1' = \mathfrak{a}_1'x + \mathfrak{c}_1'z + \mathfrak{d}_1',$$

wobei

$$\mathfrak{a}_1' = \mathfrak{a}_1, \quad \mathfrak{c}_1' = \mathfrak{c}_1 - \frac{\mathfrak{b}_1}{\mathfrak{b}_2'}\,\mathfrak{c}_2', \quad \mathfrak{d}_1' = \mathfrak{d}_1 - \frac{\mathfrak{b}_1}{\mathfrak{b}_2'}\,\mathfrak{d}_2'$$

gesetzt ist. Eine einfache Rechnung zeigt, daſs

$$\mathfrak{c}_1' = -\frac{\begin{vmatrix} \mathfrak{b}_1 & \mathfrak{c}_1 \\ \mathfrak{b}_2 & \mathfrak{c}_2 \end{vmatrix}}{\begin{vmatrix} \mathfrak{a}_1 & \mathfrak{b}_1 \\ \mathfrak{a}_2 & \mathfrak{b}_2 \end{vmatrix}}\,\mathfrak{a}_1, \quad \mathfrak{d}_1' = -\frac{\begin{vmatrix} \mathfrak{b}_1 & \mathfrak{d}_1 \\ \mathfrak{b}_2 & \mathfrak{d}_2 \end{vmatrix}}{\begin{vmatrix} \mathfrak{a}_1 & \mathfrak{b}_1 \\ \mathfrak{a}_2 & \mathfrak{b}_2 \end{vmatrix}}\,\mathfrak{a}_1$$

ist. Das Modulsystem $(\mathfrak{f}_1, \mathfrak{f}_2, \mathfrak{f}_3)$ läſst sich auf die Form $(\mathfrak{f}_1', \mathfrak{f}_2', \mathfrak{f}_3'')$ reduzieren, wo dann

$$\frac{\mathfrak{b}_2'}{\mathfrak{a}_1}\,\mathfrak{f}_1' = \begin{vmatrix} \mathfrak{a}_1 & \mathfrak{b}_1 \\ \mathfrak{a}_2 & \mathfrak{b}_2 \end{vmatrix} x + \begin{vmatrix} \mathfrak{c}_1 & \mathfrak{b}_1 \\ \mathfrak{c}_2 & \mathfrak{b}_2 \end{vmatrix} z + \begin{vmatrix} \mathfrak{d}_1 & \mathfrak{b}_1 \\ \mathfrak{d}_2 & \mathfrak{b}_2 \end{vmatrix}$$

$$\mathfrak{a}_1\,\mathfrak{f}_2' = \begin{vmatrix} \mathfrak{a}_1 & \mathfrak{b}_1 \\ \mathfrak{a}_2 & \mathfrak{b}_2 \end{vmatrix} y + \begin{vmatrix} \mathfrak{a}_1 & \mathfrak{c}_1 \\ \mathfrak{a}_2 & \mathfrak{c}_2 \end{vmatrix} z + \begin{vmatrix} \mathfrak{a}_1 & \mathfrak{d}_1 \\ \mathfrak{a}_2 & \mathfrak{d}_2 \end{vmatrix}$$

$$\mathfrak{b}_2'\,\mathfrak{f}_3'' = \begin{vmatrix} \mathfrak{a}_1 & \mathfrak{b}_1 & \mathfrak{c}_1 \\ \mathfrak{a}_2 & \mathfrak{b}_2 & \mathfrak{c}_2 \\ \mathfrak{a}_3 & \mathfrak{b}_3 & \mathfrak{c}_3 \end{vmatrix} z + \begin{vmatrix} \mathfrak{a}_1 & \mathfrak{b}_1 & \mathfrak{d}_1 \\ \mathfrak{a}_2 & \mathfrak{b}_2 & \mathfrak{d}_2 \\ \mathfrak{a}_3 & \mathfrak{b}_3 & \mathfrak{d}_3 \end{vmatrix}$$

ist. Weitere Reduktionen lassen sich vornehmen, wenn man über $\mathfrak{c}_3''$ oder über

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

eine Voraussetzung macht.

III. Wir nehmen nun an, dafs

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = 0$$

sei. Dann läfst sich das Modulsystem (f_1, f_2, f_3) auf das System (f_1', f_2', d_3'') oder — da $a_1 \neq 0$ und $b_2' \neq 0$ vorausgesetzt war — auf ein System reduzieren, das aus den folgenden beiden Funktionen und einer Konstanten

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} x + \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix} z + \begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} y + \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix} z + \begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}$$

besteht. Hieraus ergiebt sich nun sofort als hinreichende und notwendige Bedingung für die Lösbarkeit des Gleichungssystems, dafs

$$\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} = 0$$

sein mufs. Betreffs der Lösungen folgt sofort, dafs z jeden beliebigen Wert annehmen darf, dafs nachher aber x und y völlig bestimmt sind, denn es ist

$$x = -\frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z - \frac{\begin{vmatrix} d_1 & b_1 \\ d_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

$$y = -\frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} z - \frac{\begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}.$$

IV. Endlich kommen wir zu der Voraussetzung

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \neq 0,$$

die es ermöglicht, mit dem zuletzt erhaltenen Modulsystem (f_1', f_2', f_3'') noch eine weitere Reduktion vorzunehmen. Wir führen ein

$$f_1'' = f_1' - \frac{c_1'}{c_3''} f_3''$$

$$f_2'' = f_2' - \frac{c_2'}{c_3''} f_3'',$$

so dafs f_1'' und f_2'' völlig frei von z werden, f_1'' also nur von x, f_2'' nur von y, wie f_3'' nur von z abhängt, so dafs

$$f_1'' = a_1''x + d_1''$$
$$f_2'' = b_2''y + d_2''$$

ist, wo

$$a'' = a_1' = a_1, \quad d_1'' = d_1' - \frac{c_1'}{c_3''} d_3''$$

$$b_2'' = b_2', \quad d_2'' = d_2' - \frac{c_2'}{c_3''} d_3''$$

gesetzt ist. Die Berechnung der Koeffizienten kann ohne Schwierigkeiten durchgeführt werden. Man vermeidet aber einige Rechnungen, wenn man an Stelle des Systems (f_1', f_2', f_3'') das damit äquivalente System am Schlusse von II zur Reduktion benutzt. Es möge dem Leser überlassen bleiben, dies durchzuführen. Wir geben hier nur das Resultat an, dafs das Modulsystem reduziert werden kann auf ein System, in dem folgende drei Funktionen vorkommen:

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} x + \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} y + \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} z + \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix},$$

die gleich Null gesetzt — ebenso wie f_1'', f_2'', f_3'' — lehren, dafs immer nur ein einziges Lösungssystem vorhanden ist, das dargestellt werden kann in der Form

$$x = -\frac{\begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}$$

$$y = -\frac{\begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}$$

$$z = -\frac{\begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}}.$$

Fassen wir zum Schlufs die Ergebnisse der Untersuchung zusammen, so zeigt sich Folgendes: Das Modulsystem (f_1, f_2, f_3) kann sich auf vier verschiedene Arten verhalten, nämlich äquivalent sein 1) einer Konstanten; 2) einem System aus einer linearen Funktion und einer Konstanten; 3) einem System von zwei unabhängigen linearen Funktionen und einer Konstanten; 4) endlich einem System von drei unabhängigen linearen Funktionen. Diese Fälle treten beziehungsweise ein, je nachdem in der Matrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

1) alle Elemente verschwinden oder 2) wenigstens ein Element von Null verschieden ist, aber alle zu bildenden Determinanten zweiter Ordnung verschwinden oder 3) wenigstens eine von diesen sich nicht so verhält, aber die vollständige Determinante dritter Ordnung verschwindet, oder endlich 4) diese Determinante von Null verschieden ist.

Ferner zeigt sich betreffs der Lösbarkeit der Gleichungen $f_1 = 0$, $f_2 = 0$, $f_3 = 0$, dafs eine solche nur dann eintritt, wenn die in den Fällen 1) 2) 3) auftretenden Konstanten verschwinden, und dafs dann beziehungsweise 1) alle drei Unbekannten, 2) zwei Unbekannte, 3) eine einzige Unbekannte einen willkürlichen Wert haben können, worauf die übrigen völlig bestimmt sind, oder endlich 4) alle drei Unbekannten völlig bestimmte Werte haben.

Die Lösbarkeit des Systems ist aber gewährleistet stets im Falle 4), bei den übrigen Fällen aber müssen von der Matrix

$$\begin{Bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{Bmatrix}$$

1) alle Elemente verschwinden oder 2) alle zu bildenden Determinanten zweiter Ordnung oder 3) alle möglichen Determinanten dritter Ordnung.

---

§ 45. Reduktion der linearen Modulsysteme.

Da die Beantwortung der Frage nach der Auflösbarkeit eines Systems linearer Gleichungen, wie wir in § 43 und 44 gesehen haben, im Grunde genommen auf nichts anderes hinauskommt als auf die Transformation eines Modulsystems, so wollen wir uns zunächst damit beschäftigen, das lineare Modulsystem

$$\text{(A)} \qquad (f_1, f_2, \ldots f_m)$$

zu reduzieren, wo $f_1, f_2, \ldots f_m$ die linearen Funktionen

$$\text{(1)} \quad f_i = \sum_k a_{ik} x_k + a_{i0} \quad (i = 1, 2, \ldots m;\ k = 1, 2, \ldots n)$$

bedeuten sollen.

Es möge hier eine Bemerkung in § 43 wiederholt werden, dafs wir hier einen verallgemeinerten Begriff des Modulsystems zu Grunde legen und zwei Systeme von linearen Funktionen als äquivalente Modulsysteme bezeichnen, wenn sich alle Funktionen des einen Systems homogen und linear mit konstanten Koeffizienten durch die des andern Systems darstellen lassen und umgekehrt. Kommen in einem solchen Modulsystem auch Konstante vor, so kann man sie, wenn sie verschwinden, einfach weglassen; treten aber mehrere nicht verschwindende Konstanten auf, so lassen sich alle durch Multiplikation aus einer einzigen herleiten und können daher bis auf diese, die auch als Vielfaches der Einheit betrachtet werden kann, ausgeschieden werden.

Wenn alle Elemente der verkürzten Matrix verschwinden, so ist die Reduktion leicht zu vollziehen, und man erhält

$$(f_1, f_2, \dots f_m) = (a_{10}, a_{20}, \dots a_{m0}) = a,$$

wo a eine Konstante bedeutet, die entweder den Wert 0 hat, wenn die Gröfsen $a_{10}, a_{20}, \dots a_{m0}$ alle verschwinden, oder einen beliebigen von 0 verschiedenen Wert, als den man auch die Zahl 1 annehmen darf, wenn wenigstens eine dieser Gröfsen von 0 verschieden ist.

Nehmen wir nun an, dafs wenigstens ein Element der verkürzten Matrix von Null verschieden sei. Beachten wir, dafs man die Indices der Funktionen und Variabeln umändern kann, so sieht man leicht ein, dafs man ohne der Allgemeinheit Abbruch zu thun, $a_{11} \neq 0$ voraussetzen darf; wenn dies nicht der Fall ist, so kann man die Funktionen und Variabeln so bezeichnen, dafs diese Annahme zutrifft. Dann kann man das Modulsystem so umformen, dafs nur eine einzige Funktion, nämlich f_1, die Variable x_1 enthält, die übrigen Funktionen dagegen von ihr frei sind. Setzen wir nämlich

$$(2) \qquad f_2^{(1)} = f_2 - \frac{a_{21}}{a_{11}} f_1, \dots f_m^{(1)} = f_m - \frac{a_{m1}}{a_{11}} f_1,$$

so haben die Funktionen $f_2^{(1)}, \dots f_m^{(1)}$ die Form

$$(3) \qquad f_i^{(1)} = \sum_k a_{ik}^{(1)} x_k; \qquad (i = 2, 3, \dots m; \; k = 0, 2, \dots n)$$

es fehlt also in ihnen der Koeffizient von x_1. Zugleich aber ist

$$(A^{(1)})\quad (f_1, f_2, \ldots f_m) = (f_1, f_2^{(1)}, \ldots f_m^{(1)}),$$

weil die Elemente dieser beiden Modulsysteme gegenseitig linear und homogen durch einander ausdrücken lassen, wie die obigen Gleichungen (2) unmittelbar zeigen, aus denen andererseits folgt, dafs

$$(3)\qquad f_2 = f_2^{(1)} + \frac{a_{21}}{a_{11}} f_1, \ldots . f_m = f_m^{(1)} + \frac{a_{m1}}{a_{11}} f_1$$

ist.

Betrachten wir nun das Modulsystem $(f_2^{(1)}, f_3^{(1)}, \ldots f_m^{(1)})$, so können wir mit ihm genau so verfahren wie mit $(f_1, f_2, \ldots f_m)$. Sind alle Elemente der verkürzten Matrix gleich 0, so ist die Reduktion unmittelbar beendet. Ist dies nicht der Fall, so kann man die Funktionen und die Variabeln $x_2, \ldots x_n$ so bezeichnen, dafs $a_{22}^{(1)} \neq 0$ ist, und dann ein neues System bilden, in dem nur eine Funktion die Variable x_2 enthält, während in den übrigen nur $x_3, x_4, \ldots x_n$ vorkommen. Sind

$$(4)\qquad f_3^{(2)} = f_3^{(1)} - \frac{a_{32}^{(1)}}{a_{22}^{(1)}} f_2^{(1)}, \ldots . f_m^{(2)} = f_m^{(1)} - \frac{a_{m2}^{(1)}}{a_{22}^{(1)}} f_2^{(1)}$$

diese Funktionen, so ist

$$(f_2^{(1)}, f_3^{(1)}, \ldots f_m^{(1)}) = (f_2^{(1)}, f_3^{(2)}, \ldots f_m^{(2)})$$

und infolge dessen

$$(A^{(2)})\quad (f_1, f_2, f_3, \ldots f_m) = (f_1, f_2^{(1)}, f_3^{(2)}, \ldots f_m^{(2)}).$$

So kann man weiter fortfahren, bis man auf eine verkürzte Matrix mit lauter verschwindenden Elementen stöfst. Da nun aber mit jeder Transformation eine neue Variable ausgeschieden wird, so mufs man zu einem Abschlusse gelangen, und das erhaltene System

$$(A^{(r-1)})\quad (f_1, f_2^{(1)}, f_3^{(2)}, \ldots f_r^{(r-1)}, a_{r+1\,r+1}^{(r)})$$

wird Funktionen von folgender Form besitzen:

$$(5)\quad \begin{cases} f_1 = a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n + a_{10} \\ f_2^{(1)} = a_{22}^{(1)} x_2 + \ldots + a_{2n}^{(1)} x_n + a_{20}^{(1)} \\ \vdots \\ f_r^{(r-1)} = a_{rr}^{(r-1)} x_r + \ldots + a_{rn}^{(r-1)} x_n + a_{r0}^{(r-1)}, \end{cases}$$

dazu noch eine Gröſse $a^{(r)}_{r+1\,r+1}$, die man, wenn sie nicht gleich Null ist, durch jeden beliebigen davon verschiedenen Wert ersetzen darf. Von den r Variabeln $x_1, x_2, \dots x_r$ enthält f_1 jedenfalls x_1, $f^{(1)}_2$ jedenfalls x_2 u. s. w., $f^{(r-1)}_r$ jedenfalls x_r. Man kann nun das Modulsystem noch so transformieren, daſs jedes seiner Elemente von diesen Variabeln nur eine einzige enthält. $f^{(r-1)}_r$ ist schon von dieser Form, da in ihm auſser x_r nur die Variabeln $x_{r+1}, \dots x_n$ vorkommen können. Aus den andern Funktionen $f_1, f^{(1)}_2, \dots f^{(r-2)}_{r-1}$ können wir zunächst x_r fortschaffen, wenn wir

$$(6) \qquad f^{(i)}_i = f^{(i-1)}_i - \frac{a^{(i-1)}_{ir}}{a^{(r-1)}_{rr}} f^{(r-1)}_r \quad (i = 1, 2, \dots r-1)$$

einführen, wo dann $f^{(i)}_i$ die allgemeine Form hat

$$(7) \quad f^{(i)}_i = a^{(i)}_{ii} x_i + \dots + a^{(i)}_{i\,r-1} x_{r-1} + a^{(i)}_{i\,r+1} x_{r+1} + \dots + a^{(i)}_{in} x_n + a^{(i)}_{i0}, \qquad (i = 1, 2, \dots r)$$

so daſs $f^{(r-1)}_{r-1}$ die Variable x_{r-1} enthält, während auſserdem nur noch $x_{r+1}, \dots x_n$ vorkommen können.

Wir schaffen nun aus den $(r-2)$ ersten Funktionen weiter x_{r-1} fort, wenn wir setzen

$$(8) \qquad f^{(i+1)}_i = f^{(i)}_i - \frac{a^{(i)}_{i\,r-1}}{a^{(r-1)}_{r-1\,r-1}} f^{(r-1)}_{r-1}. \quad (i = 1, 2, \dots r-2)$$

So kann man weiter fortfahren, und der letzte Schritt hierbei besteht darin, daſs man aus der Funktion $f^{(r-2)}_1$ in dem System

$$(f^{(r-2)}_1, f^{(r-1)}_2, \dots f^{(r-1)}_{r-1}, f^{(r-1)}_r)$$

die Unbestimmte x_2 entfernt, indem man

$$(9) \qquad f^{(r-1)}_1 = f^{(r-2)}_1 - \frac{a^{(r-2)}_{12}}{a^{(r-1)}_{22}} f^{(r-1)}_2$$

setzt. So erhält man zum Schluſs aus dem System

$$(f_1, f^{(1)}_2, f^{(2)}_3 \dots f^{(r-2)}_{r-1}, f^{(r-1)}_r, a^{(r)}_{r+1\,r+1})$$

das äquivalente Modulsystem

$$(f_1^{(r-1)},\ f_2^{(r-1)}, \ldots f_r^{(r-1)}, a_{r+1\,r+1}^{(r)}),$$

dessen Funktionen die folgende Form haben:

$$(10)\quad \begin{cases} f_1^{(r-1)} = a_{11}^{(r-1)} x_1 + a_{1\,r+1}^{(r-1)} x_{r+1} + \ldots + a_{1n}^{(r-1)} x_n + a_{10}^{(r-1)} \\ f_2^{(r-1)} = a_{22}^{(r-1)} x_2 + a_{2\,r+1}^{(r-1)} x_{r+1} + \ldots + a_{2n}^{(r-1)} x_n + a_{20}^{(r-1)} \\ \vdots \\ f_r^{(r-1)} = a_{rr}^{(r-1)} x_r + a_{r\,r+1}^{(r-1)} x_{r+1} + \ldots + a_{rn}^{(r-1)} x_n + a_{r0}^{(r-1)}. \end{cases}$$

Hierbei ist zu bemerken, dafs durch die zuletzt erfolgte Umformung die Koeffizienten von $x_1, x_2, \ldots x_r$ nicht geändert worden sind, dafs also

$$a_{11}^{(r-1)} = a_{11},\quad a_{22}^{(r-1)} = a_{21}^{(1)}, \ldots a_{r-1\,r-1}^{(r-1)} = a_{r-1\,r-1}^{(r-2)}$$

ist.

Hat nun die mit $a_{r+1\,r+1}^{(r)}$ bezeichnete Konstante einen von Null verschiedenen Wert, so kann man noch durch Subtraktion die konstanten Glieder in den einzelnen Funktionen entfernen, indem man

$$(11)\quad f_i^{(r)} = f_i^{(r-1)} - \frac{a_{i0}^{(r-1)}}{a_{r+1\,r+1}^{(r)}} \cdot a_{r+1\,r+1}^{(r)} \qquad (i = 1, 2, \ldots r)$$

bildet. Hierdurch werden die Koeffizienten der Unbestimmten in den Funktionen

$$(12)\quad \begin{cases} f_1^{(r)} = a_{11}^{(r)} x_1 + a_{1\,r+1}^{(r)} x_{r+1} + \ldots + a_{1n}^{(r)} x_n \\ f_2^{(r)} = a_{22}^{(r)} x_2 + a_{2\,r+1}^{(r)} x_{r+1} + \ldots + a_{2n}^{(r)} x_n \\ \vdots \\ f_r^{(r)} = a_{rr}^{(r)} x_r + a_{r\,r+1}^{(r)} x_{r+1} + \ldots + a_{rn}^{(r)} x_n \end{cases}$$

gar nicht geändert.

Wir können das Resultat unserer bisherigen Betrachtungen in folgender Weise zusammenfassen:

Jedes lineare Modulsystem

$$(A)\qquad (f_1, f_2, \ldots f_m)$$

kann man entweder in das System

$$(B)\qquad (f_1^{(r-1)}, f_2^{(r-1)}, \ldots f_r^{(r-1)})$$

oder in das System

$$(B') \qquad (f_1^{(r)}, f_2^{(r)}, \ldots f_r^{(r)}, a_{r+1\,r+1}^{(r)})$$

transformieren, wo die Funktionen die unter (10) und (12) angegebene Form haben, und $a_{r+1\,r+1}^{(r)}$ eine von Null verschiedene Konstante von beliebigem Werte darstellt.

Die Form dieser Funktionen ist nun aber von besonderer Wichtigkeit für weitere Schlufsfolgerungen. Da nämlich jede eine Unbestimmte enthält, die in den übrigen nicht vorkommt, und mit einem von Null verschiedenen Koeffizienten, so kann keine linear und homogen durch die übrigen ausgedrückt werden, da in diesen allen ja eine Variable sicher nicht vorkommt, oder anders ausgedrückt, diese Funktionen sind linear unabhängig von einander. Das gilt auch noch für die Konstante $a_{r+1\,r+1}^{(r)}$ in dem Modulsystem (B'), die unabhängig von den Funktionen $f_1^{(r)}, f_2^{(r)}, \ldots f_r^{(r)}$ ist. Nun ist die Art der Transformation des Modulsystems (A) auf die Form (B) und (B') manchen Willkürlichkeiten unterworfen, so dafs man, wenn man die Funktionen und Variabeln bei der Transformation in anderer Reihenfolge heranzieht, andere Modulsysteme wie (B) und (B'), wenn auch von derselben Form, erhalten kann. Alle so erhaltenen Systeme müssen aber, da sie dem ursprünglichen System $(f_1, f_2, \ldots f_m)$ äquivalent sind, auch unter sich äquivalent sein, also müssen sich die Elemente eines Systems linear und homogen durch die des andern und umgekehrt ausdrücken lassen. Daraus folgt, dafs die Anzahl der Elemente eine unveränderliche ist. Denn betrachtet man die Elemente des einen Modulsystems als lineare Formen der des andern, so läfst sich das erstere auf ein solches reduzieren, das nicht mehr Formen als Variabeln hat, und auch das Umgekehrte findet statt. Die unveränderliche Anzahl der Elemente des reduzierten Modulsystems nennt man seinen Rang. Er hat in den beiden oben angeführten Fällen die Werte r und $r+1$, die keinenfalls die kleinere der beiden Zahlen m und $n+1$ übersteigen.

Die beiden unterschiedenen Fälle rühren nun aber davon her, dafs wir auch nichthomogene lineare Funktionen als Elemente des ursprünglichen Modulsystems zugelassen haben, wobei die von den Unbestimmten freien Koeffizienten eine

Sonderstellung beanspruchen. Hätten wir die Funktionen dadurch homogen gemacht, dafs wir die sämtlichen konstanten Koeffizienten mit dem Faktor x_0 versehen hätten, so hätte sich die Transformation genau so erledigen lassen, wie oben, nur dafs auch die von den Unbestimmten freien Koeffizienten in den Modulsystemen (B) und (B') zum Schlufs mit dem Faktor x_0 versehen werden müssen.

Daraus folgt nun, dafs der Rang des Modulsystems allein von der Matrix abhängig ist, und wir können daher dieser denselben Rang zuschreiben, wie dem aus ihnen gebildeten Modulsystem, gleichviel, ob die Elemente desselben als homogene oder nichthomogene lineare Funktionen angenommen worden sind.

Hätte man von den Funktionen des ursprünglichen Modulsystems einige weggelassen, so würde dadurch der Rang des so verkürzten Systems sicher nicht vergröfsert worden sein, aber es kann unter Umständen eintreten, dafs er sich verringert. Entfernt man aus dem Modulsystem hinreichend viele Funktionen, so tritt eine Erniedrigung seines Ranges schon aus dem Grunde ein, weil dieser niemals die Anzahl der Funktionen an Gröfse übertreffen kann. Es ist aber von nicht geringer Wichtigkeit zu bemerken, dafs man gewisse Arten der Verkürzung vornehmen kann, ohne dafs der Rang sich ändert. Schliefsen wir nämlich nicht homogene Funktionen aus, so dafs nur der Fall (B) des reduzierten Systems mit verschwindenden Koeffizienten $a_{10}^{(r-1)}, a_{20}^{(r-1)}, \ldots a_{r0}^{(r-1)}$ in den Gleichungen (10) zu betrachten ist, so ist leicht einzusehen, dafs das Modulsystem (A) oder (B) dem System

$$(f_1, f_2, \ldots f_r)$$

äquivalent ist, weil dieses allein zur Reduktion benutzt wird, vorausgesetzt, dafs keine Permutation der Indizes der Formen und Variabeln im Laufe der Transformation notwendig wird; alle übrigen Formen $f_{r+1}, f_{r+2}, \ldots f_m$ lassen sich also dann linear und homogen durch $f_1, f_2, \ldots f_r$ ausdrücken.

Der Rang kann ferner nicht gröfser werden, wenn man einige der Variabeln verschwinden läfst, denn wenn man dasselbe in dem reduzierten System thut, so bleibt die Äquivalenz zwischen den so veränderten Systemen (A) und (B) erhalten. Läfst man aber in dem System $(f_1, f_2, \ldots f_r)$

die Variabeln $x_{r+1}, x_{r+2}, \ldots x_n$ verschwinden, so ist der Rang immer noch gleich r; das reduzierte System würde dann die einfache Form:

$$(a_{11} x_1, a_{22}^{(1)} x_2, \ldots a_{rr}^{(r-1)} x_r)$$

oder:

$$(a_{11}^{(r-1)} x_1, a_{22}^{(r-1)} x_2, \ldots a_{rr}^{(r-1)} x_r)$$

erhalten haben, wo die Gröfsen $a_{ii}^{(i-1)}$ und $a_{ii}^{(r-1)}$ genau dieselben sind wie oben. Demnach ist die Matrix:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1r} \\ a_{21} & a_{22} & \ldots & a_{2r} \\ \vdots & & & \\ a_{r1} & a_{r2} & \ldots & a_{rr} \end{pmatrix}$$

genau vom Range r. Bedenken wir nun, dafs wir die Bezeichnung so eingerichtet vorausgesetzt haben, dafs keine Vertauschungen von Zeilen und Spalten mehr nötig sind, um die Transformation durchzuführen, so gelangen wir zu folgendem Satze:

Wenn eine Matrix

$$(a_{ik}) \qquad (i = 1, 2, \ldots m;\ k = 1, 2, \ldots n)$$

den Rang r hat, so haben alle quadratischen Matrizen, die man aus ihr durch Auswahl von beliebigen Zeilen und Spalten bilden kann, höchstens den Rang r, und man kann wenigstens eine einzige quadratische Matrix von r Zeilen und r Spalten

$$\begin{pmatrix} a_{i_1 k_1} & a_{i_1 k_2} & \ldots & a_{i_1 k_r} \\ a_{i_2 k_1} & a_{i_2 k_2} & \ldots & a_{i_2 k_r} \\ \vdots & & & \\ a_{i_r k_1} & a_{i_r k_2} & \ldots & a_{i_r k_r} \end{pmatrix}$$

bilden, die genau den Rang r hat. Hierbei sind $i_1, i_2, \ldots i_r$ r verschiedene Zahlen aus der Reihe $1, 2, \ldots m$, $k_1, k_2, \ldots k_r$ ebensolche aus der Reihe $1, 2, \ldots n$. Dafs die Umkehrung dieses Satzes gilt, ist unmittelbar klar und braucht nicht mehr ausführlich bewiesen zu werden. Die Aufgabe, den Rang einer Matrix zu bestimmen, ist hierdurch

zurückgeführt auf die Frage, wie man erkennen kann, ob eine quadratische Matrix von r Zeilen und Spalten den Rang r hat oder einen niederen.

---

§ 46. Nutzen der vorhergehenden Betrachtungen für die Auflösung eines Systems linearer Gleichungen.

Bevor wir unsere Untersuchungen nun weiter verfolgen, ist es angebracht, den Nutzen darzulegen, den die vorhergehenden Betrachtungen für die Auflösung eines Systems linearer Gleichungen gewähren, wobei wir an die Gleichungen (1) und (10) des vorigen Paragraphen anknüpfen.

Da die Funktionen des ursprünglichen Systems $(f_1, f_2, \ldots f_m)$ linear und homogen mit denen des reduzierten Systems $(f_1^{(r-1)}, f_2^{(r-1)}, \ldots f_r^{(r-1)}, a_{r+1\,r+1}^{(r)})$ zusammenhängen, so ist ihr Verschwinden wechselseitig bedingt, d. h. das Gleichungssystem:

$$f_1 = 0,\ f_2 = 0, \ldots f_m = 0$$

ist vollständig gleichbedeutend mit dem Gleichungssystem:

$$f_1^{(r-1)} = 0,\ f_2^{(r-1)} = 0, \ldots f_{r-1}^{(r-1)} = 0,$$
$$f_r^{(r-1)} = 0,\ a_{r+1\,r+1}^{(r)} = 0.$$

Alle Werte, die das erste System befriedigen, genügen auch dem letzteren und umgekehrt. Wir brauchen uns also nur an das letztere zu halten, das uns unmittelbar die Lösungen liefert in der Form:

$$x_1 = -\frac{a_{1\,r+1}^{(r-1)} x_{r+1} + a_{1\,r+2}^{(r-1)} x_{r+2} + \ldots + a_{1n}^{(r-1)} x_n + a_{10}^{(r-1)}}{a_{11}^{(r-1)}}$$

$$x_2 = -\frac{a_{2\,r+1}^{(r-1)} x_{r+1} + a_{2\,r+2}^{(r-1)} x_{r+2} + \ldots + a_{2n}^{(r-1)} x_n + a_{20}^{(r-1)}}{a_{22}^{(r-1)}}$$

$$\vdots$$

$$x_r = -\frac{a_{r\,r+1}^{(r-1)} x_{r+1} + a_{r\,r+2}^{(r-1)} x_{r+2} + \ldots + a_{rn}^{(r-1)} x_n + a_{r0}^{(r-1)}}{a_{rr}^{(r-1)}},$$

wobei $x_{r+1}, x_{r+2}, \ldots x_n$ ganz beliebige Werte haben können. Da die Anzahl dieser Unbestimmten gleich $(n - r)$ ist, so pflegt man das Ergebnis kurz so auszudrücken:

Ein System linearer Gleichungen mit n Unbekannten und vom Range r hat eine $(n - r)$fache Mannigfaltigkeit von Lösungen oder gar keine.

Der letzte Fall tritt nämlich ein, wenn die bei der Reduktion des Modulsystems auftretende Konstante $a^{(r)}_{r+1\,r+1} \neq 0$ ist. Besteht das Gleichungssystem aus lauter homogenen linearen Funktionen, so tritt die Konstante gar nicht auf. Ein solches System besitzt also stets Lösungen, und wenn der Rang r seinen höchsten Wert $r = n$ erreicht, so wird das System nur dann befriedigt, wenn man sämtlichen Unbekannten den Wert Null giebt. Sind die Funktionen dagegen nicht homogen, so ist es von Bedeutung, die Bedingung für die Lösbarkeit einfacher zu formulieren:

Es mufs der Rang der verkürzten Matrix

$$\begin{pmatrix} a_{11}\ a_{12} \dots a_{1n} \\ \vdots \\ a_{m1}\ a_{m2} \dots a_{mn} \end{pmatrix}$$

übereinstimmen mit dem der vollständigen Matrix

$$\begin{pmatrix} a_{10}\ a_{11} \dots a_{1n} \\ \vdots \\ a_{m0}\ a_{m1} \dots a_{mn} \end{pmatrix}.$$

---

§ 47. Quadratische Matrizen und Determinanten.

Wie wir in § 45 gesehen haben, kommt die Frage, von welchem Range eine Matrix ist, darauf hinaus, die hinreichenden und notwendigen Bedingungen dafür zu ermitteln, dafs eine quadratische Matrix von n Zeilen und n Spalten genau den Rang n hat.

Wir nehmen an, wie es nach früheren Bemerkungen gestattet ist, dafs diese Matrix

$$\begin{pmatrix} a_{11}\ a_{12} \dots a_{1n} \\ a_{21}\ a_{22} \dots a_{2n} \\ \vdots \\ a_{n1}\ a_{n2} \dots a_{nn} \end{pmatrix}$$

schon so angeordnet ist, dafs keine Vertauschungen von Zeilen oder Spalten mehr nötig sind, um die in § 45 besprochene Transformation zu vollziehen, und richten nun unsere Aufmerksamkeit auf die Formen der entstehenden neuen Matrizen.

Die erste Transformation führt zu der Matrix:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & & & \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix},$$

wobei allgemein

$$a_{ik}^{(1)} = a_{ik} - \frac{a_{i1}\,a_{1k}}{a_{11}} \qquad (i, k = 2, 3 \dots n)$$

ist, die zweite zu:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \dots & a_{3n}^{(2)} \\ \vdots & & & & \\ 0 & 0 & a_{n3}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix},$$

wo

$$a_{ik}^{(2)} = a_{ik}^{(1)} - \frac{a_{i2}^{(1)}\;a_{2k}^{(1)}}{a_{22}^{(1)}}. \qquad (i, k = 3, 4 \dots n)$$

Allgemein hat die h te Transformation die Matrix:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1\,h+1} & \dots & a_{1n} \\ 0 & a_{21}^{(1)} & \dots & a_{2\,h+1}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & & & & & \\ 0 & 0 & \dots & a_{h+1\,h+1}^{(h)} & \dots & a_{h+1\,n}^{(h)} \\ \vdots & & & & & \\ 0 & 0 & \dots & a_{n\,h+1}^{(h)} & \dots & a_{nn}^{(h)} \end{pmatrix}$$

zur Folge, und es ist

$$a_{ik}^{(h)} = a_{ik}^{(h-1)} - \frac{a_{ih}^{(h-1)}\,a_{hk}^{(h-1)}}{a_{hh}^{(h-1)}}. \qquad (i, k = h+1,\ h+2, \dots n)$$

Zuletzt, nach $(n - 1)$ maliger Transformation, gelangt man zu der Matrix

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1\,n-1} & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2\,n-1}^{(1)} & a_{2n}^{(1)} \\ \vdots & & & & \\ 0 & 0 & \dots & a_{n-1\,n-1}^{(n-2)} & a_{n-1\,n}^{(n-2)} \\ 0 & 0 & \dots & 0 & a_{nn}^{(n-1)} \end{pmatrix},$$

wobei

$$a_{nn}^{(n-1)} = a_{nn}^{(n-2)} - \frac{a_{nn-1}^{(n-2)} a_{n-1n}^{(n-2)}}{a_{n-1n-1}^{(n-2)}}$$

ist. Diese Matrix entspricht einem Modulsystem, das wir nach Früherem mit

$$(f_1, f_2^{(1)}, f_3^{(2)}, \ldots f_{n-1}^{(n-2)}, f_n^{(n-1)})$$

zu bezeichnen haben.

Aus den hingeschriebenen Transformationsformeln für die neu entstehenden Elemente ersieht man, dafs sich jedes neue Element der i-ten Zeile stets linear und homogen durch Elemente der i-ten Zeile der vorhergegangenen Matrix darstellt. Daraus folgt, dafs es auch eine lineare und homogene Funktion der i-ten Zeile in der ursprünglichen Matrix, also der Gröfsen $a_{i1}, a_{i2}, \ldots a_{in}$, ist. Insbesondere ist also $a_{nn}^{(n-1)}$ homogen und linear durch $a_{n1}, a_{n2} \ldots a_{nn}$ darstellbar.

Soll nun die zuletzt erhaltene Matrix vom Range n sein, so ist dazu notwendig, dafs die Gröfsen $a_{11}, a_{22}^{(1)}, \ldots a_{nn}^{(n-1)}$ sämtlich von Null verschieden sind, und das ist der Fall, wenn das Produkt $a_{11}\, a_{22}^{(1)}\, a_{33}^{(2)} \ldots a_{nn}^{(n-1)}$ diese Eigenschaft hat. Dieses Produkt läfst sich nun als Funktion der n-Elemente a_{ik} der ursprünglichen Matrix darstellen und soll als solches durch das Zeichen

$$\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

oder noch kürzer durch

$$|a_{ik}| \qquad (i, k = 1, 2, \ldots n)$$

bezeichnet und Determinante der Matrix (a_{ik}) $(i, k = 1, 2, \ldots n)$ genannt werden.

Für die einfachsten Fälle $n = 2$ und $n = 3$ ist es leicht, die Determinante als Funktion der Elemente des ursprünglichen Matrix hinzuschreiben.

Für $n = 2$ ergiebt sich aus

$$a_{22}^{(1)} = a_{22} - \frac{a_{21}\, a_{12}}{a_{11}}$$

durch Multiplikation mit a_{11}

$$a_{11}\, a_{22}^{(1)} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\, a_{22} - a_{12}\, a_{21}.$$

Für $n = 3$ hat man in den Ausdruck

$$a_{33}^{(2)} = a_{33}^{(1)} - \frac{a_{32}^{(1)}\, a_{23}^{(1)}}{a_{22}^{(1)}}$$

die Werte

$$a_{22}^{(1)} = a_{22} - \frac{a_{21}\, a_{12}}{a_{11}}, \quad a_{23}^{(1)} = a_{23} - \frac{a_{21}\, a_{13}}{a_{11}}$$

$$a_{32}^{(1)} = a_{32} - \frac{a_{31}\, a_{12}}{a_{11}}, \quad a_{33}^{(1)} = a_{33} - \frac{a_{31}\, a_{13}}{a_{11}}$$

einzusetzen. Eine leichte Rechnung, bei der sich zwei Glieder fortheben, zeigt, dafs

$$a_{11}\, a_{22}^{(1)}\, a_{33}^{(2)} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{matrix} = a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{13}\, a_{21}\, a_{32} \\ - a_{13}\, a_{22}\, a_{31} - a_{12}\, a_{21}\, a_{33} - a_{11}\, a_{23}\, a_{32} \end{matrix}$$

ist.

So könnte man fortfahren in der Berechnung der Fälle $n = 4, 5 \ldots$ Aber man gelangt auf diese Weise zu immer verwickelteren Ausdrücken, ohne das Gesetz zu erkennen, das ihrer Bildung zu Grunde liegt. Wir müssen daher zu neuen Betrachtungen übergehen, und hierbei bietet sich naturgemäfs als Ausgangspunkt die Bemerkung dar, dafs die Determinante ihrer Definition nach durchaus abhängig ist von der Art und Weise, wie wir die Zeilen und Spalten der Matrix angeordnet haben. Daraus ergiebt sich also die Aufgabe, zu untersuchen, welchen Änderungen die Determinante unterliegt, wenn man Zeilen und Spalten unter sich vertauscht. Bevor wir jedoch diese Frage allgemein untersuchen, wollen wir erst zusehen, wie sich die Sache für die Fälle $n = 2$ und $n = 3$ verhält.

Für den Fall $n = 2$ erkennt man durch Ausrechnung, dafs die Determinante ihr Vorzeichen wechselt, wenn man die Zeilen oder die Spalten mit einander vertauscht. Es ist nämlich

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix} = a_{21}\, a_{12} - a_{11}\, a_{22} = -(a_{11}\, a_{22} - a_{21}\, a_{12}) = - \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

$$\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix} = a_{12} a_{21} - a_{22} a_{11} = -(a_{11} a_{22} - a_{21} a_{12}) = - \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

aber wieder

$$\begin{vmatrix} a_{22} & a_{21} \\ a_{12} & a_{11} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

wie man aufser durch Ausrechnung auch noch daraus erkennt, dafs die Determinanten auseinander durch Vertauschung der Zeilen und dann der Spalten entstehen.

Betrachten wir im Falle $n = 3$ zunächst eine beliebige Permutation der Zeilen. Wie wir früher gesehen haben, lassen sich alle 6 Permutationen von drei Elementen aus einer cyklischen Permutation von ihnen und einer beliebigen Transposition zusammensetzen. Was die Wirkung der ersteren betrifft, so ergiebt sich leicht aus

$$\begin{vmatrix} a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{vmatrix} = \begin{matrix} a_{31}\, a_{12}\, a_{23} + a_{32}\, a_{13}\, a_{21} + a_{33}\, a_{11}\, a_{22} \\ - a_{33}\, a_{12}\, a_{21} - a_{32}\, a_{11}\, a_{23} - a_{31}\, a_{13}\, a_{22}, \end{matrix}$$

dafs sich die Determinante nicht ändert, dagegen ergiebt sich bei Transposition der ersten beiden Zeilen aus

$$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{21}\, a_{12}\, a_{33} + a_{22}\, a_{13}\, a_{31} + a_{23}\, a_{11}\, a_{32} \\ - a_{23}\, a_{12}\, a_{31} - a_{22}\, a_{11}\, a_{33} - a_{21}\, a_{13}\, a_{32}, \end{matrix}$$

dafs die Determinante ihr Vorzeichen wechselt. Da nun aus dieser Transposition und der cyklischen Permutation jede beliebige Vertauschung der Zeilen sich bilden läfst, so ist unsere Frage bezüglich der Zeilenvertauschungen völlig gelöst, und wir können folgenden Satz aussprechen: Die Determinante ändert ihr Vorzeichen bei jeder Vertauschung zweier Zeilen unter einander. Für die Spalten aber gelten durchaus die analogen Entwickelungen, wovon sich der Leser selber überzeugen mag.

Wir schreiten jetzt zum allgemeinen Fall und beantworten die Frage, wie sich die Determinante ändert, wenn man ihre Zeilen beliebigen Permutation unterwirft. Da eine solche sich aus lauter Transpositionen zusammensetzen läfst, so wollen wir allgemein untersuchen, was geschieht, wenn die $h$te und die $(h+1)$te Zeile mit einander vertauscht werden.

Zunächst ist ersichtlich, dafs dann die Reduktion der Matrix in Bezug auf die (h — 1) ersten Zeilen genau dieselbe bleibt, wie vorhin, dafs sich also insbesondere die Gröfsen a_{11}, $a_{22}^{(1)} \ldots$, $a_{h-1\,h-1}^{(h-2)}$ gar nicht ändern, weil sie nur von den Elementen der (h — 1) ersten Zeilen abhängig sind. Eine wirkliche Änderung tritt erst von der hten Zeile ab ein. So lange man noch nicht die hte Transformation vollzogen hat, besteht diese einfach darin, dafs die hte und (h + 1)Zeile vertauscht werden. Um nun die Veränderungen nach der hten und den folgenden Transformationen beurteilen zu können, drücken wir die durch sie entstehenden Elemente durch die Elemente der (h — 1)ten transformierten Matrix aus; aber aus der Gleichung

$$a_{hh}^{(h-1)}\, a_{h+1\,h+1}^{(h)} = \begin{vmatrix} a_{hh}^{(h-1)} & a_{h\,h+1}^{(h-1)} \\ a_{h+1\,h}^{(h-1)} & a_{h+1\,h+1}^{(h-1)} \end{vmatrix}$$

erkennen wir, dafs $a_{hh}^{(h-1)}\, a_{h+1\,h+1}^{(h)}$ einen Vorzeichenwechsel erleidet, wenn die hte und (h + 1) Zeile mit einander vertauscht werden. Da aber weiter, wie man sich durch Rechnung überzeugt,

$$a_{hh}^{(h-1)}\, a_{h+1\,h+1}^{(h)}\, a_{ik}^{(h+1)} = \begin{vmatrix} a_{hh}^{(h-1)} & a_{h\,h+1}^{(h-1)} & a_{hk}^{(h-1)} \\ a_{h+1\,h}^{(h-1)} & a_{h+1\,h+1}^{(h-1)} & a_{h+1\,k}^{(h-1)} \\ a_{ih}^{(h-1)} & a_{i\,h+1}^{(h-1)} & a_{ik}^{(h-1)} \end{vmatrix}$$

$$(i, k = h + 2, \ldots n)$$

ist, so ergiebt sich dasselbe für $a_{hh}^{(h-1)}\, a_{h+1\,h+1}^{(h)}\, a_{ik}^{(h+1)}$ aus den Eigenschaften der dreigliedrigen Determinanten. Daher wird $a_{ik}^{(h+1)}$ überhaupt gar nicht geändert. Setzen wir nun, nachdem dies festgestellt ist, in die Gleichung

$$a_{ik}^{(t+1)} = a_{ik}^{(t)} - \frac{a_{i\,t+1}^{(t)}\, a_{t+1\,k}^{(t)}}{a_{t+1\,t+1}^{(t)}}$$

für t der Reihe nach h + 1, h + 2, …, n — 1 ein, so erkennen wir weiter, dafs auch die Gröfsen $a_{ik}^{(h+2)}$, $a_{ik}^{(h+3)}$, … $a_{nn}^{(n-1)}$ bei der Vertauschung ihren Wert behalten. Als Resultat der Betrachtungen ergiebt sich also, dafs das Produkt $a_{11}\, a_{22}^{(1)} \ldots a_{nn}^{(n-1)}$ oder die Determinante bei der Vertauschung zweier aufeinander folgender Zeilen, der hten

und (h + 1)ten Zeile, der Matrix einen Vorzeichenwechsel erleidet, ihren absoluten Wert dagegen beibehält. Wir haben nun früher nicht nur gezeigt, dafs sich aus solchen Vertauschungen durch wiederholte Anwendung alle möglichen Permutationen herleiten lassen, sondern auch eine Einteilung derselben in zwei Klassen kennen gelernt, von denen die der ersten sich durch eine gerade, der der zweiten durch eine ungerade Anzahl von Transpositionen darstellen lassen (§ 40). Mit Beziehung hierauf können wir den allgemeinen Satz aussprechen:

Eine Determinante von n^2 Elementen a_{ik} $(i, k = 1, 2, \ldots n)$ behält ihren Wert bei, wenn man ihre Zeilen den $\frac{n!}{2}$ Permutationen der Alterngruppe unterwirft, wird dagegen durch alle andern $\frac{n!}{2}$ Permutationen der Zeilen nur im Vorzeichen, nicht im absoluten Betrage geändert.

§ 48. Darstellung der Determinanten.

Nach den Vorbereitungen im vorigen Paragraphen ist es nun nicht schwer, die Determinante $|a_{ik}|$ $(i, k = 1, \ldots n)$ als Funktion der Elemente a_{ik} darzustellen. Wir haben schon bemerkt, dafs $a_{nn}^{(n-1)}$ eine lineare Form der n Gröfsen $a_{n1}, a_{n2}, \ldots a_{nn}$ ist, während $a_{11}, a_{22}^{(1)}, \ldots a_{n-1\,n-1}^{(n-2)}$ von diesen unabhängig sind. Da die Determinante das Produkt dieser Gröfsen ist, so ist sie selbst auch eine lineare Form der n Elemente der nten Zeile. Nun haben wir aber gesehen, dafs sie nur ihr Vorzeichen ändert, wenn man an Stelle von $a_{n1}, a_{n2}, \ldots a_{nn}$ die Gröfsen $a_{h1}, a_{h2}, \ldots a_{hn}$ setzt, wo h eine der Zahlen $1, 2, \ldots n-1$ bedeuten kann, und hieraus ergiebt sich dann sofort, dafs die Determinante eine lineare Form der Elemente einer jeden Zeile ist.

Daher mufs es möglich sein, die Determinante in folgender Form darzustellen

$$\underset{(i, k = 1, \ldots n)}{|a_{ik}|} = \sum_{i_1, i_2, \ldots i_n} \varepsilon_{i_1 i_2 \ldots i_n} a_{1 i_1} a_{2 i_2} \ldots a_{n i_n}, \quad (i_1, i_2, \ldots i_n = 1, 2, \ldots n)$$

wo die Gröfsen $\varepsilon_{i_1 i_2 \ldots i_n}$ von den Elementen a_{ik} unabhängige Koeffizienten sind, deren Beziehungen zu einander sich sofort angeben lassen. Vertauscht man nämlich in der Determinante irgend zwei Zeilen, etwa die hte und kte mit einander, so tritt nur eine Vorzeichenänderung ein; es gilt demnach die Gleichung

$$\sum_{(i_1, i_2, \ldots i_n)} \varepsilon_{i_1 \ldots i_h \ldots i_k \ldots i_n} a_{1 i_1} \ldots a_{h i_h} \ldots a_{k i_k} \ldots a_{n i_n}$$
$$= - \sum_{(i_1, i_2, \ldots i_n)} \varepsilon_{i_1 \ldots i_k \ldots i_h \ldots i_n} a_{1 i_1} \ldots a_{h i_h} \ldots a_{k i_k} \ldots a_{n i_n}.$$

Da diese eine identische ist, so müssen die entsprechenden Koeffizienten auf beiden Seiten gleich sein; es ist also

$$\varepsilon_{i_1 \ldots i_h \ldots i_k \ldots i_n} = - \varepsilon_{i_1 \ldots i_k \ldots i_h \ldots i_n}.$$

Der Koeffizient $\varepsilon_{i_1 i_2 \ldots i_n}$ ändert also sein Vorzeichen, aber nicht seinen Betrag, wenn zwei Indices mit einander vertauscht werden. Sind nun zwei Indizes gleich, so ist daher der Koeffizient gleich Null. Wir brauchen also nur solche Koeffizienten zu betrachten, in denen $i_1, i_2 \ldots i_n$ verschieden sind, also in irgend einer Reihenfolge mit den Zahlen $1, 2, \ldots n$ übereinstimmen und sich daher aus dieser Zahlenreihe durch eine bestimmte Permutation ableiten lassen. Nun ist aber leicht erkennbar, dafs der Koeffizient $\varepsilon_{12\ldots n}$ in der Determinante den Wert $+1$ besitzt. Daraus ergeben sich nun sofort folgende Sätze:

1) Der Koeffizient $\varepsilon_{i_1 i_2 \ldots i_n}$ von $a_{1 i_1} a_{2 i_2} \ldots a_{n i_n}$ ist $+1$ oder -1, je nachdem die Permutation

$$\begin{pmatrix} 1 & 2 \ldots n \\ i_1 & i_2 \ldots i_n \end{pmatrix}$$

durch eine gerade oder ungerade Anzahl von Transpositionen darstellbar ist, also der Alterngruppe angehört oder nicht.

2) Die Anzahl der von Null verschiedenen Glieder einer Determinante von n^2 Elementen ist gleich $P(n) = n!$ Unter diesen sind ebensoviele positive wie negative vorhanden, also von jeder Art $\frac{n!}{2}$.

3) Jedes Glied enthält ein und nur ein Element aus jeder Zeile und ebenso eins und nur eins aus jeder Spalte.

Hieraus folgt, dafs sich die Determinante auch in der Form

$$\sum_{(h_1, h_2 \dots h_n)} \eta_{h_1 h_2 \dots h_n} a_{h_1 1} a_{h_2 2} \dots a_{h_n n}$$

darstellen lassen mufs, wo die Gröfsen $\eta_{h_1 h_2 \dots h_n}$ die Werte $+1$ oder -1 haben und sich auf folgende Weise bestimmen lassen. Ordnen wir das Produkt $a_{h_1 1} a_{h_2 2} \dots a_{h_n n}$ so, dafs die ersten Indizes $h_1, h_2 \dots h_n$ in der Reihenfolge $1, 2, \dots n$ erscheinen, und sind dann die zweiten Indizes der Reihe nach $i_1, i_2 \dots i_n$, so ist $\eta_{h_1 h_2 \dots h_n} = \varepsilon_{i_1 i_2 \dots i_n}$.

Die Permutationen

$$\begin{pmatrix} 1 & 2 \dots n \\ h_1 & h_2 \dots h_n \end{pmatrix} \text{ und } \begin{pmatrix} 1 & 2 \dots n \\ i_1 & i_2 \dots i_n \end{pmatrix}$$

sind nun aber invers zu einander, daher gehören beide entweder der Alterngruppe zugleich an oder nicht; es ist also stets $\eta_{h_1 h_2 \dots h_n} = \varepsilon_{i_1 i_2 \dots i_n}$, und wir können die Determinante auch in der Form schreiben

$$\sum_{(i_1, i_2 \dots i_n)} \varepsilon_{i_1 i_2 \dots i_n} a_{i_1 1} a_{i_2 2} \dots a_{i_n n}.$$

Hieraus folgt der Satz:

4) Eine Determinante bleibt ungeändert, wenn die Zeilen zu Spalten und die Spalten zu Zeilen gemacht werden. (§ 11.)

---

§ 49. Zusammenstellung der Haupteigenschaften der Determinanten.

Nachdem wir die Gleichberechtigung der Zeilen und Spalten in den Determinanten erkannt haben, stellen wir jetzt die Haupteigenschaften, die bisher teilweise nur für die Zeilen als richtig nachgewiesen waren, in verallgemeinerter Form zusammen:

I. Die Determinante ist eine lineare Form der Elemente einer beliebigen Reihe (Zeile oder Spalte).

Aus dieser Eigenschaft ergeben sich sofort folgende Sätze (nach § 8 I und II):

1. Man kann eine Determinante mit einer Zahl multiplizieren oder durch eine Zahl dividieren, indem man die sämtlichen Elemente einer beliebigen Reihe mit der Zahl multipliziert oder durch sie dividiert.

Hiernach ist

$$\begin{vmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ \vdots & & \\ a_{i-1\,1} & a_{i-1\,2} & \ldots a_{i-1\,n} \\ a_{i1}t & a_{i2}t & \ldots a_{in}t \\ a_{i+1\,1} & a_{i+1\,2} & \ldots a_{i+1\,n} \\ \vdots & & \\ a_{n1} & a_{n2} & \ldots a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ \vdots & & \\ a_{i-1\,1} & a_{i-1\,2} & \ldots a_{i-1\,n} \\ a_{i1} & a_{i2} & \ldots a_{in} \\ a_{i+1\,1} & a_{i+1\,2} & \ldots a_{i+1\,n} \\ \vdots & & \\ a_{n1} & a_{n2} & \ldots a_{nn} \end{vmatrix} t,$$

$$\begin{vmatrix} a_{11} \ldots a_{1k-1} & a_{1k}t & a_{1k+1} \ldots a_{1n} \\ a_{21} \ldots a_{2k-1} & a_{2k}t & a_{2k+1} \ldots a_{2n} \\ \vdots & & \\ a_{n1} \ldots a_{nk-1} & a_{nk}t & a_{nk+1} \ldots a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} \ldots a_{1k-1} & a_{1k} & a_{1k+1} \ldots a_{1n} \\ a_{21} \ldots a_{2k-1} & a_{2k} & a_{2k+1} \ldots a_{2n} \\ \vdots & & \\ a_{n1} \ldots a_{nk-1} & a_{nk} & a_{nk+1} \ldots a_{nn} \end{vmatrix} t.$$

2. Sind die Elemente einer Reihe in einer Determinante Aggregate aus gleichviel Gliedern, so ist die Determinante gleich der Summe derjenigen Determinanten, die man erhält, wenn man an Stelle der Summen die einzelnen Glieder setzt.

Also ist

$$\begin{vmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ \vdots & & \\ a_{i-1\,1} & a_{i-1\,2} & \ldots a_{i-1\,n} \\ \sum\limits_{h=1}^{h=m} a_{i1}^{(h)} & \sum\limits_{h=1}^{h=m} a_{i2}^{(h)} & \ldots \sum\limits_{h=1}^{h=m} a_{in}^{(h)} \\ a_{i+1\,1} & a_{i+1\,2} & \ldots a_{i+1\,n} \\ \vdots & & \\ a_{n1} & a_{n2} & \ldots a_{nn} \end{vmatrix} = \sum_{h=1}^{h=m} \begin{vmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ \vdots & & \\ a_{i-1\,1} & a_{i-1\,2} & \ldots a_{i-1\,n} \\ a_{i1}^{(h)} & a_{i2}^{(h)} & \ldots a_{in}^{(h)} \\ a_{i+1\,1} & a_{i+1\,2} & \ldots a_{i+1\,n} \\ \vdots & & \\ a_{n1} & a_{n2} & \ldots a_{nn} \end{vmatrix}$$

und ebenso

$$\begin{vmatrix} a_{11} \ldots a_{1\,k-1} \sum\limits_{h=1}^{h=m} a_{1k}^{(h)} a_{1\,k+1} \ldots a_{1n} \\ a_{21} \ldots a_{2\,k-1} \sum\limits_{h=1}^{h=m} a_{2k}^{(h)} a_{2\,k+1} \ldots a_{2n} \\ \vdots \\ a_{n1} \ldots a_{n\,k-1} \sum\limits_{h=1}^{h=m} a_{nk}^{(h)} a_{n\,k+1} \ldots a_{nn} \end{vmatrix} = \sum_{h=1}^{h=m} \begin{vmatrix} a_{11} \ldots a_{1\,k-1} a_{1k}^{(h)} a_{1k+1} \ldots a_{1n} \\ a_{21} \ldots a_{2\,k-1} a_{2k}^{(h)} a_{2k+1} \ldots a_{2n} \\ \vdots \\ a_{n1} \ldots a_{n\,k-1} a_{nk}^{(h)} a_{nk+1} \ldots a_{nn} \end{vmatrix}.$$

II. Die Determinante wechselt ihr Vorzeichen, bleibt aber dem absoluten Werte nach ungeändert, wenn zwei Parallelreihen mit einander vertauscht werden.

Aus dieser zweiten Haupteigenschaft leitet man ab:

3. Eine Determinante mit entsprechend gleichen Elementen in zwei Parallelreihen hat den Wert Null.

Denn dann kann man die beiden Parallelreihen mit einander vertauschen, ohne dafs sie ihren Wert ändert. Andererseits ändert sie aber ihr Vorzeichen. Beides kann vereint nur geschehen, wenn der Wert verschwindet. (Aus $a = -a$ folgt $2a = 0$, $a = 0$.)

4. Permutirt man in einer Determinante Zeilen und Spalten, so bleibt sie ungeändert oder wechselt das Vorzeichen, je nachdem die beiden Permutationen komponiert der Alterngruppe angehören oder nicht.

Es ist also

$$\begin{vmatrix} a_{i_1 1} & a_{i_1 2} & \ldots & a_{i_1 n} \\ a_{i_2 1} & a_{i_2 2} & \ldots & a_{i_2 n} \\ \vdots \\ a_{i_n 1} & a_{i_n 2} & \ldots & a_{i_n n} \end{vmatrix} = \varepsilon_{i_1 i_2 \ldots i_n} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

$$\begin{vmatrix} a_{1k_1} & a_{1k_2} & \ldots & a_{1k_n} \\ a_{2k_1} & a_{2k_2} & \ldots & a_{2k_n} \\ \vdots \\ a_{nk_1} & a_{nk_2} & \ldots & a_{nk_n} \end{vmatrix} = \varepsilon_{k_1 k_2 \ldots k_n} \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix},$$

$$\begin{vmatrix} a_{i_1 k_1} & a_{i_1 k_2} & \dots & a_{i_1 k_n} \\ a_{i_2 k_1} & a_{i_2 k_2} & \dots & a_{i_2 k_n} \\ \vdots \\ a_{i_n k_1} & a_{i_n k_2} & \dots & a_{i_n k_n} \end{vmatrix} = \varepsilon_{i_1 i_2 \dots i_n} \; \varepsilon_{k_1 k_2 \dots k_n} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Aus den beiden Haupteigenschaften I und II ergiebt sich endlich:

5. Der Wert einer Determinante bleibt ungeändert, wenn man die Elementen einer Reihe um Vielfache der entsprechenden Elemente einer parallelen Reihe vermehrt oder vermindert.

Denn es ist z. B. nach (2)

$$\begin{vmatrix} a_{11} + a_{12} t & a_{12} & \dots & a_{1n} \\ a_{21} + a_{22} t & a_{22} & \dots & a_{2n} \\ \vdots \\ a_{n1} + a_{n2} t & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{12} t & a_{12} & \dots & a_{1n} \\ a_{22} t & a_{22} & \dots & a_{2n} \\ \vdots \\ a_{n2} t & a_{n2} & \dots & a_{nn} \end{vmatrix};$$

hierbei läfst sich nach (1) die zweite Determinante umformen in das Produkt

$$\begin{vmatrix} a_{12} & a_{12} & \dots & a_{1n} \\ a_{22} & a_{22} & \dots & a_{2n} \\ \vdots \\ a_{n2} & a_{n2} & \dots & a_{nn} \end{vmatrix} t,$$

dessen erster Faktor nach (3) verschwindet.

---

§ 50. Unterdeterminanten.

Jede Determinante ist, wie wir bewiesen haben, eine lineare Form der Elemente einer beliebigen Reihe, deren Koeffizienten nur von den Elementen der anderen Reihen abhängig sind. Diese Koeffizienten wollen wir jetzt genauer betrachten und allgemein den Koeffizienten von a_{ik} in der Determinante

$$A = | a_{ik} | \qquad (i, k = 1, 2, \dots n)$$

mit A_{ki} bezeichnen, so dafs mit der früher (§ 6) auseinandergesetzten Bezeichnungsweise der Differentialrechnung auch

$$A_{ki} = \frac{\delta A}{\delta a_{ik}}$$

gesetzt werden kann. Bevor wir auf die Form dieser sogenannten zu a_{ik} adjungierten Gröfsen A_{ki} und die Art ihrer Abhängigkeit von den Elementen der Determinante näher eingehen, wollen wir die Gleichungen betrachten, denen sie Genüge leisten.

I. Entwickeln wir die Determinante nach den Elementen der i ten Zeile, so ergiebt sich

$$A = \sum_h a_{kh} A_{hk}, \qquad (h, k = 1, 2, \ldots n)$$

aus der Entwickelung nach der k ten Spalte dagegen folgt

$$A = \sum_h a_{hk} A_{kh}. \qquad (h, k = 1, 2, \ldots n).$$

Ersetzt man nun in der ersteren Relation die Elemente a_{kh} durch a_{ih}, so entsteht aus A eine Determinante, in der die i te und k te Zeile entsprechend gleiche Elemente haben, die also verschwinden mufs. Da das Analoge auch für die Spalten gilt, so ergeben sich für $i \neq k$ die beiden Gleichungssysteme

$$\begin{aligned} \sum_h a_{ih} A_{hk} &= 0 \\ \sum_h a_{hi} A_{kh} &= 0. \end{aligned} \qquad (h = 1, 2, \ldots n).$$

Beide Arten von Relationen kann man zusammenfassen, wenn man mit dem Zeichen δ_{ik} die Zahl 1 oder 0 bezeichnet, je nachdem $i = k$ oder $i \neq k$ ist, in der Form

$$\begin{aligned} \sum_h a_{ih} A_{hk} &= \delta_{ik} A \\ \sum_h a_{hi} A_{kh} &= \delta_{ik} A. \end{aligned} \qquad (h = 1, 2, \ldots n)$$

Mit Hülfe des Begriffs der Komposition der Matrizen kann man ferner die sämtlichen Relationen in einer einzigen Formel vereinigen. Wir betrachten neben den n^2 Gröfsen a_{ik} die n^2 Koeffizienten A_{ik}, die zu den ersteren das adjungierte System bilden.

Dann ist

$$\begin{Bmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots & \\ a_{n1} & a_{n2} \dots a_{nn} \end{Bmatrix} \begin{Bmatrix} A_{11} & A_{12} \dots A_{1n} \\ A_{21} & A_{22} \dots A_{2n} \\ \vdots & \\ A_{n1} & A_{n2} \dots A_{nn} \end{Bmatrix} = \begin{Bmatrix} A & 0 & 0 \dots & 0 \\ 0 & A & 0 \dots & 0 \\ \vdots & & & \\ 0 & 0 & 0 \dots & A \end{Bmatrix}$$

oder in noch kürzerer Form

$$(a_{ik})\ (A_{ik}) = (\delta_{ik}\, A). \qquad (i, k = 1, \dots n)$$

II. Wir wenden uns jetzt zur Bestimmung der Form der adjungierten Elemente A_{ik}, von denen wir zeigen wollen, dafs sie wieder Determinanten sind.

Am einfachsten gestaltet sich die Untersuchung, wenn wir zunächst A_{11} bestimmen und dann daraus allgemein A_{ik} ableiten. Da die Determinante A durch die Summe

$$\sum_{(i_1 i_2 \dots i_n)} \varepsilon_{i_1 i_2 \dots i_n}\, a_{1 i_1}\, a_{2 i_2} \dots a_{n i_n}$$

dargestellt wird, so ergiebt sich für A_{11} als Koeffizient von a_{11} die Summe:

$$A_{11} = \sum_{(i_2, i_3 \dots i_n)} \varepsilon_{1 i_2 \dots i_n}\, a_{2 i_2}\, a_{3 i_3} \dots a_{n i_n}.$$

Diese ist aber, wie unmittelbar ersichtlich ist, gleich der Determinante

$$\begin{vmatrix} a_{22} & a_{23} \dots a_{2n} \\ a_{32} & a_{33} \dots a_{3n} \\ \vdots & \\ a_{n2} & a_{n3} \dots a_{nn} \end{vmatrix},$$

die also einfach daraus aus der Determinante A hervorgeht, dafs man aus ihr die erste Zeile und erste Spalte fortläfst.

Um nun A_{ki} allgemein zu bestimmen, ändern wir die Determinante durch Permutation der Zeilen und Spalten, so um, dafs von den Zeilen die i te die erste wird, die übrigen dann in ihrer früheren Ordnung aufeinanderfolgen, dafs ferner von den Spalten die k te die erste wird, während die übrigen ihre frühere Reihenfolge beibehalten, so dafs also aus A die Determinante

$$A = (-1)^{i+k} \begin{vmatrix} a_{ik} & a_{i1} & \dots a_{i\,k-1} & a_{i\,k+1} & \dots a_{in} \\ a_{1k} & a_{11} & \dots a_{1\,k-1} & a_{1\,k+1} & \dots a_{1n} \\ \vdots \\ a_{i-1\,k} & a_{i-1\,1} & \dots a_{i-1\,k-1} & a_{i-1\,k+1} & \dots a_{i-1\,n} \\ a_{i+1\,k} & a_{i+1\,1} & \dots a_{i+1\,k-1} & a_{i+1\,k+1} & \dots a_{i+1\,n} \\ \vdots \\ a_{nk} & a_{n1} & \dots a_{n\,k-1} & a_{n\,k+1} & \dots a_{nn} \end{vmatrix}$$

wird. In der so eingerichteten Determinante ergiebt sich nun als Koeffizient von a_{ik} unmittelbar

$$A_{ki} = (-1)^{i+k} \begin{vmatrix} a_{11} & \dots a_{1\,k-1} & a_{1\,k+1} & \dots a_{1n} \\ \vdots \\ a_{i-1\,1} & \dots a_{i-1\,k-1} & a_{i-1\,k+1} & \dots a_{i-1\,n} \\ a_{i+1\,1} & \dots a_{i+1\,k-1} & a_{i+1\,k+1} & \dots a_{i+1\,n} \\ \vdots \\ a_{n1} & \dots a_{n\,k-1} & a_{n\,k+1} & \dots a_{nn} \end{vmatrix}.$$

Man erhält also allgemein den Koeffizienten A_{ki} von a_{ik} aus der Determinante A dadurch, dafs man aus ihr die $i$te Zeile und $k$te Spalte wegläfst und der so entstehenden Determinante das Vorzeichen $(-1)^{i+k}$ beifügt. Da die Gröfsen A_{ik} Determinanten sind, die $(n-1)^2$ Elemente haben, so werden sie auch als Unterdeterminanten $(n-1)$ter Ordnung der Determinante A bezeichnet. Von diesen kann man wieder die Unterdeterminanten bilden und so fortfahren bis man zuletzt auf die n^2 Unterdeterminanten erster Ordnung a_{ik} stöfst. Allgemein giebt es n_i^2 Unterdeterminanten $i$ter Ordnung zu A, wo n_i (wie in § 5) den Binomialkoeffizienten $\frac{n(n-1)\dots(n-i+1)}{2.3\dots i}$ bedeutet.

III. Wir wollen von der Entwickelung einer Determinanten nach den Elementen einer Reihe nun noch eine Anwendung geben.

Betrachten wir die Determinante

$$\begin{vmatrix} a_{00} & a_{01} \dots a_{0n} \\ a_{10} & a_{11} \dots a_{1n} \\ \vdots \\ a_{n0} & a_{n1} \dots a_{nn} \end{vmatrix},$$

so ergiebt sich, wenn man nach den Elementen der ersten Zeile entwickelt,

$$a_{00}\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + \sum_k (-1)^k a_{0k} \begin{vmatrix} a_{10} \dots a_{1k-1} & a_{1k+1} \dots a_{1n} \\ a_{20} \dots a_{2k-1} & a_{2k+1} \dots a_{2n} \\ \vdots & \\ a_{n0} \dots a_{nk-1} & a_{nk+1} \dots a_{nn} \end{vmatrix}.$$

$(k = 1, 2, \dots n)$

Ebenso leitet man ab, dafs

$$\begin{vmatrix} a_{10} \dots a_{1k-1} & a_{1k+1} \dots a_{1n} \\ a_{20} \quad a_{2k-1} & a_{2k+1} \dots a_{2n} \\ \vdots & \\ a_{n0} \dots a_{nk-1} & a_{nk+1} \dots a_{nn} \end{vmatrix}$$

$$= \sum_i (-1)^{i-1} a_{i0} \begin{vmatrix} a_{11} & \dots a_{1k-1} & a_{1k+1} & \dots a_{1n} \\ \vdots & & & \\ a_{i-11} & \dots a_{i-1k-1} & a_{i-1k+1} & \dots a_{i-1n} \\ a_{i+11} & a_{i+1k-1} & a_{i+1k+1} & \dots a_{i+1n} \\ \vdots & & & \\ a_{n1} & \dots a_{nk-1} & a_{nk+1} & \dots a_{nn} \end{vmatrix}$$

$$= \sum_i (-1)^{k-1} A_{ki}\, a_{i0}, \qquad (i = 1, 2, \dots n)$$

wo A_{ki} die in II betrachteten adjungierten Elemente der Determinante

$$A = |a_{ik}| \qquad (i, k = 1, 2, \dots n)$$

sind. Setzt man alles ein, so folgt

$$\begin{vmatrix} a_{00} & a_{01} \dots a_{0n} \\ a_{10} & a_{11} \dots a_{1n} \\ \vdots & \\ a_{n0} & a_{n1} \dots a_{nn} \end{vmatrix} = A a_{00} - \sum_{i,k} A_{ki}\, a_{0k}\, a_{i0}. \qquad (i, k = 1, 2, \dots n)$$

---

§ 51. Berechnung von Determinanten.

I. Die in § 48 gegebene analytische Darstellung der Determinanten giebt das einfachste Mittel an die Hand, eine gegebene Determinante zu berechnen. Da aber die Bildung der Determinanten höherer Ordnung ein wenig umständlich ist, so empfiehlt es sich, die Determinanten nach den Ele-

menten einer beliebigen Reihe zu entwickeln und die als Koeffizienten auftretenden adjungierten Unterdeterminanten für sich gesondert zu berechnen. Die folgenden Beispiele werden danach dem Leser ohne weitere Erläuterungen verständlich sein.

1) $$\begin{vmatrix} x_1 - x & x_2 - x \\ y_1 - y & y_2 - y \end{vmatrix} = (x_1 - x)(y_2 - y) - (y_1 - y)(x_2 - x)$$
$$= x_1 y_2 - y_1 x_2 + x_2 y - y_2 x + x y_1 - y x_1.$$

2) $$\begin{vmatrix} \cos\alpha & \sin\alpha \\ \sin\beta & \cos\beta \end{vmatrix} = \cos\alpha\cos\beta - \sin\alpha\sin\beta = \cos(\alpha + \beta).$$

3) $$\begin{vmatrix} a - x & b \\ c & d - x \end{vmatrix} = (a - x)(d - x) - bc$$
$$= x^2 - (a + d)x + ad - bc.$$

4) $$\begin{vmatrix} ax^2 + 2bx + c & ax + b \\ ax + b & a \end{vmatrix} = a(ax^2 + 2bx + c)$$
$$- (ax + b)^2 = ac - b^2.$$

5) $$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_2 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$
$$= a_1(b_2 c_3 - c_2 b_3) - b_1(a_2 c_3 - a_3 c_2) + c_1(a_2 b_3 - a_3 b_2)$$
$$= a_1 b_2 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_3 b_1 c_2 + a_2 b_3 c_1 - a_3 b_2 c_1.$$

Vgl. § 44.

6) $$\begin{vmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{vmatrix} = 2abc.$$

7) $$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} = abc - af^2 - bg^2 - ch^2 + 2fgh.$$

8) $$\begin{vmatrix} a_1 - x & b_1 & c_1 \\ a_2 & b_2 - x & c_2 \\ a_3 & b_3 & c_3 - x \end{vmatrix} = -x^3 + (a_1 + b_2 + c_3)x^2$$
$$- \left[\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} + \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} + \begin{vmatrix} c_3 & a_3 \\ c_1 & a_1 \end{vmatrix} \right] x + \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

$$9)\quad \begin{vmatrix} 1 & \cos\gamma & \cos\beta \\ \cos\gamma & 1 & \cos\alpha \\ \cos\beta & \cos\alpha & 1 \end{vmatrix} = 1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma$$

$$= 4\sin\frac{\alpha+\beta+\gamma}{2}\sin\frac{-\alpha+\beta+\gamma}{2}\sin\frac{\alpha-\beta+\gamma}{2}\sin\frac{\alpha+\beta-\gamma}{2}.$$

$$10)\quad \begin{vmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & 0 & 0 \\ a_{32} & a_{33} & 0 \\ a_{42} & a_{43} & a_{44} \end{vmatrix}$$

$$= a_{11}\, a_{22}\begin{vmatrix} a_{33} & 0 \\ a_{43} & a_{44} \end{vmatrix} = a_{11}\, a_{22}\, a_{31}\, a_{44}.$$

$$11)\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 0 \\ b_{11} & b_{12} & b_{13} & 1 & 0 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 1 & 0 \\ b_{31} & b_{32} & b_{33} & 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 \\ b_{11} & b_{12} & b_{13} & 1 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & a_{13} & 0 \\ a_{21} & a_{22} & a_{23} & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ b_{11} & b_{12} & b_{13} & 1 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

$$12)\quad \begin{vmatrix} 1 & 0 & 0 & \alpha \\ 0 & 1 & 0 & \beta \\ 0 & 0 & 1 & \gamma \\ \alpha & \beta & \gamma & 1 \end{vmatrix} = 1 - (\alpha^2 + \beta^2 + \gamma^2).$$

$$13)\quad \begin{vmatrix} 1 & c & b & \alpha \\ c & 1 & a & \beta \\ b & a & 1 & \gamma \\ \alpha & \beta & \gamma & 1 \end{vmatrix} = 1 - (\alpha^2 + \beta^2 + \gamma^2) - (a^2 + b^2 + c^2)$$

$$+ (a^2\alpha^2 + b^2\beta^2 + c^2\gamma^2) + 2(a\beta\gamma + b\gamma\alpha + c\alpha\beta) - 2(bc\beta\gamma + ca\gamma\alpha + ab\alpha\beta) + 2abc.$$

14) Die Matrix

$$\begin{vmatrix} 1 & 0 & 0 & \alpha & \alpha' \\ 0 & 1 & 0 & \beta & \beta' \\ 0 & 0 & 1 & \gamma & \gamma' \\ \alpha & \beta & \gamma & 1 & \varepsilon \\ \alpha' & \beta' & \gamma' & \varepsilon & 1 \end{vmatrix}$$

ist mindestens vom Range 3, da die Determinante 3. Grades in ihr

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1$$

ist. Es soll nachgewiesen werden, dafs, wenn sie von keinem höheren Range ist, die Relationen bestehen

$$\alpha^2 + \beta^2 + \gamma^2 = 1, \qquad \alpha'^2 + \beta'^2 + \gamma'^2 = 1$$
$$\varepsilon = \alpha\alpha' + \beta\beta' + \gamma\gamma'$$

und umgekehrt.

II. Bei den folgenden Beispielen ist von dem Satze Gebrauch gemacht, dafs man in jeder Determinante die Elemente einer Reihe um Vielfache der entsprechenden Elemente einer parallelen Reihe vermindern oder vermehren kann, ohne den Wert der Determinante zu ändern. Gelingt es auf diese Weise alle Elemente einer bis auf ein einziges zum Verschwinden zu bringen, so ergiebt sich durch Entwickelung der Determinante nach den Elementen dieser Reihe eine Reduktion auf eine Determinante niederer Ordnung. Sind die Elemente der Determinante ganze Zahlen, so kann die Determinante auf diese Weise vollständig berechnet werden.

$$15) \quad \begin{vmatrix} 7 & 8 & 3 & 2 \\ 5 & 4 & 2 & 1 \\ 0 & 2 & 3 & 5 \\ 4 & 7 & 3 & 2 \end{vmatrix} = \begin{vmatrix} -3 & 0 & -1 & 0 \\ 5 & 4 & 2 & 1 \\ -25 & -18 & -7 & 0 \\ -6 & -1 & -1 & 0 \end{vmatrix}$$

$$= \begin{vmatrix} -3 & 0 & -1 \\ -25 & -18 & -7 \\ -6 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 0 & 0 & -1 \\ -4 & -18 & -7 \\ -3 & -1 & -1 \end{vmatrix} = - \begin{vmatrix} -4 & -18 \\ -3 & -1 \end{vmatrix}$$
$$= -(4 - 54) = +50.$$

Von den Elementen der ersten, dritten und vierten Zeile wurde resp. das 2-, 5-, 2-fache der entsprechenden Elemente der zweiten Zeile subtrahiert. Nachdem dann die Determinante auf eine solche dritter Ordnung zurückgeführt war, ergab sich durch Subtraktion der 3fachen Elemente der dritten Spalte von den entsprechenden der ersten eine Reduktion auf eine Determinante zweiter Ordnung, die leicht berechnet wird. Das Verfahren läfst der Willkür grofsen Spielraum.

$$16)\quad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & 0 \\ 0 & -2 & 0 & -2 \\ 0 & 0 & -2 & -2 \end{vmatrix}$$

$$= \begin{vmatrix} -2 & -2 & 0 \\ -2 & 0 & -2 \\ 0 & -2 & -2 \end{vmatrix} = \begin{vmatrix} -2 & -2 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & -2 \end{vmatrix} = (-2)\begin{vmatrix} 2 & -2 \\ -2 & -2 \end{vmatrix}$$

$$= (-2)(-4-4) = 16.$$

Die erste Zeile ist beibehalten, und ihre Elemente sind von denen der übrigen subtrahiert worden, wodurch dann leicht eine Determinante dritter Ordnung hervorgeht. In dieser ist dann wieder die erste Zeile zur Reduktion der zweiten benutzt.

$$17)\quad \begin{vmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{vmatrix} = \begin{vmatrix} 1 & x & y \\ 0 & x_1 - x & y_1 - y \\ 0 & x_2 - x & y_2 - y \end{vmatrix} = \begin{vmatrix} x_1 - x & y_1 - y \\ x_2 - x & y_2 - y \end{vmatrix}.$$

$$18)\quad \begin{vmatrix} 1 & x_1 - x & y_1 - y \\ 1 & x_2 - x & y_2 - y \\ 1 & x_3 - x & y_3 - y \end{vmatrix} = \begin{vmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{vmatrix}.$$

Vergleiche Beispiel 1.

$$19)\quad \begin{vmatrix} a_1 x + b_1 y + c_1 z & a_1 & b_1 & c_1 \\ a_2 x + b_2 y + c_2 z & a_2 & b_2 & c_2 \\ a_3 x + b_3 y + c_3 z & a_3 & b_3 & c_3 \\ a_4 x + b_4 y + c_4 z & a_4 & b_4 & c_4 \end{vmatrix} = 0.$$

III. Bei den folgenden Beispielen ist aufserdem noch vom Satze Anwendung gemacht, dafs man gemeinschaftliche Faktoren der Elemente einer Reihe herausheben kann.

$$20)\quad \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^2 \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 \\ 0 & x_3 - x_1 & x_3^2 - x_1^2 \end{vmatrix}$$

$$= \begin{vmatrix} x_2 - x_1 & (x_2 - x_1)(x_2 + x_1) \\ x_3 - x_1 & (x_3 - x_1)(x_3 + x_1) \end{vmatrix} = (x_2 - x_1)(x_3 - x_1)\begin{vmatrix} 1 & x_1 + x_2 \\ 1 & x_1 + x_3 \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

$$21)\quad \begin{vmatrix} 1 & x_1 & x_1^3 \\ 1 & x_2 & x_2^3 \\ 1 & x_3 & x_3^3 \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^3 \\ 0 & x_2 - x_1 & x_2^3 - x_1^3 \\ 0 & x_3 - x_1 & x_3^2 - x_1^3 \end{vmatrix}$$

$$= \begin{vmatrix} x_2 - x_1 & (x_2 - x_1)(x_1^2 + x_1 x_2 + x_2^2) \\ x_3 - x_1 & (x_3 - x_1)(x_1^2 + x_1 x_3 + x_3^2) \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & x_1^2 + x_1 x_2 + x_2^2 \\ 1 & x_1^2 + x_1 x_3 + x_3^2 \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1) \begin{vmatrix} 1 & x_1^2 + x_1 x_2 + x_2^2 \\ 0 & (x_3 - x_2)(x_1 + x_2 + x_3) \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1)(x_3 - x_2)(x_1 + x_2 + x_3).$$

$$22)\quad \begin{vmatrix} 1 & x_1 + a_0 & x_1^2 + b_1 x_1 + b_0 \\ 1 & x_2 + a_0 & x_2^2 + b_1 x_2 + b_0 \\ 1 & x_3 + a_0 & x_3^2 + b_1 x_3 + b_0 \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^2 + b_1 x_1 \\ 1 & x_2 & x_2^2 + b_1 x_2 \\ 1 & x_3 & x_3^2 + b_1 x_3 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_1 - x_2)(x_2 - x_3)(x_3 - x_1).$$

23) Diese Determinantengleichung wollen wir verallgemeinern und zur Abkürzung setzen

$$f_0(x) = 1,\quad f_1(x) = x + a_{10},\quad f_2(x) = x^2 + a_{21} x + a_{20}, \ldots$$
$$\ldots f_i(x) = x^i + a_{i\,i-1} x^{i-1} + \ldots + a_{i0}, \ldots$$
$$f_{n-1}(x) = x^{n-1} + a_{n-1\,n-2} x^{n-2} + \ldots + a_{n-1\,0}.$$

Betrachten wir davon die Determinante

$$|f_{i-1}(x_k)|, \qquad (i, k = 1, 2, \ldots n)$$

so erhalten wir zunächst:

$$\begin{vmatrix} 1 & x_1 & f_2(x_1) \ldots f_{n-1}(x_1) \\ 1 & x_2 & f_2(x_2) \ldots f_{n-1}(x_2) \\ \vdots & & \\ 1 & x_n & f_2(x_n) \ldots f_{n-1}(x_n) \end{vmatrix},$$

daraus:

$$\begin{vmatrix} 1 & x_1 & x_1^2 & f_3(x_1) \ldots f_{n-1}(x_1) \\ 1 & x_2 & x_2^2 & f_3(x_2) \ldots f_{n-1}(x_2) \\ \vdots & \vdots & & \\ 1 & x_n & x_n^2 & f_3(x_n) \ldots f_{n-1}(x_n) \end{vmatrix}.$$

Da das Verfahren fortgesetzt werden kann, so ergiebt sich zunächst als Resultat:

$$\begin{vmatrix} 1 & x_1 & x_1^2 \dots x_1^{n-1} \\ 1 & x_2 & x_2^2 \dots x_2^{n-1} \\ \vdots \\ 1 & x_n & x_n^2 \dots x_n^{n-1} \end{vmatrix}.$$

24) Die zuletzt erhaltene Determinante, die auch kurz in der Form

$$\left| x_i^{k-1} \right| \qquad (i, k = 1, 2, \dots n)$$

geschrieben werden kann, ist eine neue Verallgemeinerung des Beispiels 20) und kann in folgender Weise weiter behandelt werden:

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots x_1^{n-1} \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & \dots x_2^{n-1} - x_1^{n-1} \\ \vdots \\ 0 & x_n - x_1 & x_n^2 - x_1^2 & \dots x_n^{n-1} - x_1^{n-1} \end{vmatrix}$$

$$= (x_2 - x_1)(x_3 - x_1) \dots (x_n - x_1) \times$$

$$\begin{vmatrix} 1 & x_1 + x_2 & x_1^2 + x_1 x_2 + x_2^2 \dots x_1^{n-2} + x_1^{n-3} x_2 + \dots + x_2^{n-2} \\ 1 & x_1 + x_3 & x_1^2 + x_1 x_3 + x_3^2 \dots x_1^{n-2} + x_1^{n-3} x_3 + \dots + x_3^{n-2} \\ \vdots \\ 1 & x_1 + x_n & x_1^2 + x_1 x_n + x_n^2 \dots x_1^{n-2} + x_1^{n-3} x_n + \dots + x_n^{n-2} \end{vmatrix}.$$

Wendet man nun auf die erhaltene Determinante dieselbe Art der Umformung wie beim vorigen Beispiel an, so ergiebt sie sich in der Form:

$$\begin{vmatrix} 1 & x_2 & x_2^2 \dots x_2^{n-2} \\ 1 & x_3 & x_3^2 \dots x_3^{n-2} \\ \vdots \\ 1 & x_n & x_2^2 \dots x_n^{n-2} \end{vmatrix},$$

die mit der ursprünglichen Determinante übereinstimmt, nur dafs n durch n — 1 ersetzt ist. Fährt man mit dieser in der Untersuchung fort, so ergiebt sich als Schlufsresultat:

$$\left| x_i^{k-1} \right| = \prod_{g,h} (x_g - x_h). \qquad (g, h, i, k = 1, 2, \dots n; \; g > h)$$

Setzt man

$$F(x) = (x - x_1)(x - x_2)\dots(x - x_n),$$

so kann man den Wert des Quadrates der Determinante auch in der Form

$$(-1)^{\frac{n(n-1)}{2}} \prod_{i=1}^{i=n} F'(x_i)$$

darstellen. (Vgl. § 6.)

$$25)\quad \begin{vmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{vmatrix} = \begin{vmatrix} a+b+c+d & b & c & d \\ a+b+c+d & a & d & c \\ a+b+c+d & d & a & b \\ a+b+c+d & c & b & a \end{vmatrix}$$

$$= (a+b+c+d) \begin{vmatrix} 1 & b & c & d \\ 1 & a & d & c \\ 1 & d & a & b \\ 1 & c & b & a \end{vmatrix}$$

$$= (a+b+c+d) \begin{vmatrix} 1 & b & c & d \\ 0 & a-b & d-c & c-d \\ 0 & d-b & a-c & b-d \\ 0 & c-b & b-c & a-d \end{vmatrix}$$

$$= (a+b+c+d) \begin{vmatrix} a-b & d-c & c-d \\ d-b & a-c & b-d \\ c-b & b-c & a-d \end{vmatrix}$$

$$= (a+b+c+d) \begin{vmatrix} a-b & d-c & 0 \\ d-b & a-c & a+b-c-d \\ c-b & b-c & a+b-c-d \end{vmatrix}$$

$$= (a+b+c+d)(a+b-c-d) \begin{vmatrix} a-b & d-c & 0 \\ d-b & a-c & 1 \\ c-b & b-c & 1 \end{vmatrix}$$

$$= (a+b+c+d)(a+b-c-d) \begin{vmatrix} a-b & d-c & 0 \\ d-c & a-b & 0 \\ c-b & b-c & 1 \end{vmatrix}$$

$$= (a+b+c+d)(a+b-c-d) \begin{vmatrix} a-b & d-c \\ d-c & a-b \end{vmatrix}$$

$$= (a+b+c+d)(a+b-c-d)\,[(a-b)^2 - (d-c)^2]$$

$$= (a+b+c+d)(a+b-c-d)(a-b+d-c) \times (a-b-d+c).$$

$$26)\quad \begin{vmatrix} 0 & a^2 & b^2 & c^2 \\ a^2 & 0 & z^2 & y^2 \\ b^2 & z^2 & 0 & x^2 \\ c^2 & y^2 & x^2 & 0 \end{vmatrix} = -(ax+by+cz)\times$$

$$(-ax+by+cz)(ax-by+cz)(ax+by-cz).$$

$$27)\quad \begin{vmatrix} \frac{1}{\alpha-\lambda} & \frac{1}{\alpha-\mu} & \frac{1}{\alpha-\nu} \\ \frac{1}{\beta-\lambda} & \frac{1}{\beta-\mu} & \frac{1}{\beta-\nu} \\ \frac{1}{\gamma-\lambda} & \frac{1}{\gamma-\mu} & \frac{1}{\gamma-\nu} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{1}{\alpha-\lambda} & \frac{1}{\alpha-\mu} & \frac{1}{\alpha-\nu} \\ \frac{\alpha-\beta}{(\alpha-\lambda)(\beta-\lambda)} & \frac{\alpha-\beta}{(\alpha-\mu)(\beta-\mu)} & \frac{\alpha-\beta}{(\alpha-\nu)(\beta-\nu)} \\ \frac{\alpha-\gamma}{(\alpha-\lambda)(\gamma-\lambda)} & \frac{\alpha-\gamma}{(\alpha-\mu)(\gamma-\mu)} & \frac{\alpha-\gamma}{(\alpha-\nu)(\gamma-\nu)} \end{vmatrix}$$

$$= \frac{(\alpha-\beta)(\alpha-\gamma)}{(\alpha-\lambda)(\alpha-\mu)(\alpha-\nu)} \begin{vmatrix} 1 & 1 & 1 \\ \frac{1}{\beta-\lambda} & \frac{1}{\beta-\mu} & \frac{1}{\beta-\nu} \\ \frac{1}{\gamma-\lambda} & \frac{1}{\gamma-\mu} & \frac{1}{\gamma-\nu} \end{vmatrix}$$

$$= \frac{(\alpha-\beta)(\alpha-\gamma)}{(\alpha-\lambda)(\alpha-\mu)(\alpha-\nu)} \begin{vmatrix} 1 & 0 & 0 \\ \frac{1}{\beta-\lambda} & \frac{\mu-\lambda}{(\beta-\lambda)(\beta-\mu)} & \frac{\nu-\lambda}{(\beta-\lambda)(\beta-\nu)} \\ \frac{1}{\gamma-\lambda} & \frac{\mu-\lambda}{(\gamma-\lambda)(\gamma-\mu)} & \frac{\nu-\lambda}{(\gamma-\lambda)(\gamma-\nu)} \end{vmatrix}$$

$$= \frac{(\alpha-\beta)(\alpha-\gamma)}{(\alpha-\lambda)(\alpha-\mu)(\alpha-\nu)} \, \frac{(\mu-\lambda)(\nu-\lambda)}{(\beta-\lambda)(\gamma-\lambda)} \begin{vmatrix} \frac{1}{\beta-\mu} & \frac{1}{\beta-\nu} \\ \frac{1}{\gamma-\mu} & \frac{1}{\gamma-\nu} \end{vmatrix}$$

$$= \frac{(\alpha-\beta)(\alpha-\gamma)}{(\alpha-\lambda)(\beta-\lambda)(\gamma-\lambda)} \, \frac{(\mu-\lambda)(\nu-\lambda)}{(\alpha-\mu)(\alpha-\nu)} \begin{vmatrix} \frac{1}{\beta-\mu} & \frac{\mu-\nu}{(\beta-\mu)(\beta-\nu)} \\ \frac{1}{\gamma-\mu} & \frac{\mu-\nu}{(\gamma-\mu)(\gamma-\nu)} \end{vmatrix}$$

$$=\frac{(\alpha-\beta)(\alpha-\gamma)(\beta-\gamma)(\mu-\lambda)(\lambda-\nu)(\mu-\nu)}{(\alpha-\lambda)(\beta-\lambda)(\gamma-\lambda)(\alpha-\mu)(\beta-\mu)(\gamma-\mu)(\alpha-\nu)(\beta-\nu)(\gamma-\nu)}.$$

28) Dieses Beispiel wollen wir verallgemeinern und die Determinante

$$\left|\frac{1}{\alpha_i-t_k}\right| \qquad (i, k=1, 2, \dots n)$$

betrachten. Man erhält

$$\begin{vmatrix} \frac{1}{\alpha_1-t_1} & \frac{1}{\alpha_1-t_2} & \dots & \frac{1}{\alpha_1-t_n} \\ \frac{1}{\alpha_2-t_1} & \frac{1}{\alpha_2-t_2} & \dots & \frac{1}{\alpha_2-t_n} \\ \vdots & & & \\ \frac{1}{\alpha_n-t_1} & \frac{1}{\alpha_n-t_2} & \dots & \frac{1}{\alpha_n-t_n} \end{vmatrix}$$

$$=\begin{vmatrix} \frac{1}{\alpha_1-t_1} & \frac{1}{\alpha_1-t_2} & \dots & \frac{1}{\alpha_1-t_n} \\ \frac{\alpha_1-\alpha_2}{(\alpha_1-t_1)(\alpha_2-t_1)} & \frac{\alpha_1-\alpha_2}{(\alpha_1-t_2)(\alpha_2-t_2)} & \dots & \frac{\alpha_1-\alpha_2}{(\alpha_1-t_n)(\alpha_2-t_n)} \\ \vdots & & & \\ \frac{\alpha_1-\alpha_n}{(\alpha_1-t_1)(\alpha_n-t_1)} & \frac{\alpha_1-\alpha_n}{(\alpha_1-t_2)(\alpha_n-t_2)} & \dots & \frac{\alpha_1-\alpha_n}{(\alpha_1-t_n)(\alpha_n-t_n)} \end{vmatrix}$$

$$=\frac{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)\dots(\alpha_1-\alpha_n)}{(\alpha_1-t_1)(\alpha_1-t_2)\dots(\alpha_1-t_n)}\begin{vmatrix} 1 & 1 & \dots & 1 \\ \frac{1}{\alpha_2-t_1} & \frac{1}{\alpha_2-t_2} & \dots & \frac{1}{\alpha_2-t_n} \\ \vdots & & & \\ \frac{1}{\alpha_n-t_1} & \frac{1}{\alpha_n-t_2} & \dots & \frac{1}{\alpha_n-t_n} \end{vmatrix},$$

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ \frac{1}{\alpha_2-t_1} & \frac{1}{\alpha_2-t_2} & \dots & \frac{1}{\alpha_2-t_n} \\ \vdots & & & \\ \frac{1}{\alpha_n-t_1} & \frac{1}{\alpha_n-t_2} & \dots & \frac{1}{\alpha_n-t_n} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & \cdots & 0 \\ \frac{1}{\alpha_2 - t_1} & \frac{t_2 - t_1}{(\alpha_2 - t_1)(\alpha_2 - t_2)} & \cdots & \frac{t_n - t_1}{(\alpha_2 - t_1)(\alpha_2 - t_n)} \\ \vdots & & & \\ \frac{1}{\alpha_n - t_1} & \frac{t_2 - t_1}{(\alpha_n - t_1)(\alpha_n - t_2)} & \cdots & \frac{t_n - t_1}{(\alpha_n - t_1)(\alpha_n - t_n)} \end{vmatrix}$$

$$= \frac{(t_2 - t_1)(t_3 - t_1)\ldots(t_n - t_1)}{(\alpha_2 - t_1)(\alpha_3 - t_1)\ldots(\alpha_n - t_1)} \begin{vmatrix} \frac{1}{\alpha_2 - t_2} & \cdots & \frac{1}{\alpha_2 - t_n} \\ \vdots & & \\ \frac{1}{\alpha_n - t_2} & \cdots & \frac{1}{\alpha_n - t_n} \end{vmatrix}$$

und damit folgende Reduktionsformel

$$\begin{vmatrix} \frac{1}{\alpha_1 - t_1} & \frac{1}{\alpha_1 - t_2} & \cdots & \frac{1}{\alpha_1 - t_n} \\ \frac{1}{\alpha_2 - t_1} & \frac{1}{\alpha_2 - t_2} & \cdots & \frac{1}{\alpha_2 - t_n} \\ \vdots & & & \\ \frac{1}{\alpha_n - t_1} & \frac{1}{\alpha_n - t_2} & \cdots & \frac{1}{\alpha_n - t_n} \end{vmatrix}$$

$$= \frac{1}{\alpha_1 - t_1} \cdot \frac{(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)\ldots(\alpha_1 - \alpha_n)}{(\alpha_1 - t_2)(\alpha_1 - t_3)\ldots(\alpha_1 - t_n)} \times$$

$$\frac{(t_2 - t_1)(t_3 - t_1)\ldots(t_n - t_1)}{(\alpha_2 - t_1)(\alpha_3 - t_1)\ldots(\alpha_n - t_1)} \begin{vmatrix} \frac{1}{\alpha_2 - t_2} & \cdots & \frac{1}{\alpha_2 - t_n} \\ \vdots & & \\ \frac{1}{\alpha_n - t_2} & \cdots & \frac{1}{\alpha_n - t_n} \end{vmatrix},$$

durch deren mehrmalige Anwendung sich als Wert der Determinante der folgende Ausdruck ergiebt:

$$\frac{\prod\limits_{i,k}(\alpha_i - \alpha_k) \prod\limits_{i,k}(t_k - t_i)}{\prod\limits_{i,k}(\alpha_i - t_k) \prod\limits_{i,k}(\alpha_k - t_i) \prod\limits_{i}(\alpha_i - t_i)} \quad (i, k = 1, 2, \ldots n;\ i < k)$$

$$= \frac{\prod\limits_{i,k}(\alpha_i - \alpha_k) \prod\limits_{i,k}(t_k - t_i)}{\prod\limits_{g,h}(\alpha_g - t_h)}, \quad (g, h, i, k = 1, 2, \ldots n;\ i < k)$$

dessen Quadrat man bei Einführung der beiden Funktionen

$$F(t) = (t - t_1)(t - t_2) \dots (t - t_n)$$
$$R(t) = (\alpha_1 - t)(\alpha_2 - t) \dots (\alpha_n - t)$$

auch die Form

$$\prod_i \frac{F'(t_i) R'(\alpha_i)}{F(\alpha_i)} = \prod_i \frac{F'(t_i) R'(\alpha_i)}{R(t_i)} \qquad (i = 1, 2, \dots n)$$

geben kann. (Vgl. § 6 und oben Beispiel 24.)

---

§ 52. Form der reduzierten linearen Modulsysteme.

Wir haben früher die Reduktion der linearen Modulsysteme ohne die Zuhülfenahme des Determinantenbegriffes durchgeführt. Nachdem wir jetzt die Eigenschaften der Determinanten kennen gelernt haben, ist es leicht, die dort gewonnenen Resultate wieder abzuleiten und in Rücksicht auf ihre Form zu vervollkommnen.

Wir nehmen an, dafs die verkürzte Matrix des Modulsystems

$$(f_1, f_2, \dots f_m)$$

den Rang r hat, es selbst also den Rang r oder $r + 1$ hat, und dafs die Bezeichnung der Funktionen und Variabeln so gewählt ist, dafs die Determinante

$$A = |a_{ik}| \qquad (i, k = 1, 2, \dots r)$$

einen von Null verschiedenen Wert hat. Wie wir früher gesehen haben, ist das Modulsystem dann dem folgenden äquivalent

$$(f_1, f_2, \dots f_r),$$

abgesehen von einer Konstanten, die noch etwa hinzuzufügen wäre. Mit diesem müssen wir uns daher vorläufig allein beschäftigen. Bilden wir die Determinante

$$\begin{vmatrix} a_{11} \dots a_{1k-1} \; f_1 \; a_{1k+1} \dots a_{1r} \\ a_{21} \dots a_{2k-1} \; f_2 \; a_{2k+1} \dots a_{2r} \\ \vdots \\ a_{r1} \dots a_{rk-1} \; f_r \; a_{rk+1} \dots a_{rr} \end{vmatrix},$$

so ist diese linear durch $f_1, \dots f_r$ darstellbar in der Form

$$\sum_h A_{kh}\, f_h, \qquad (h = 1, 2, \dots r)$$

wenn wir mit A_{ki} die zu a_{ik} adjungierten Elemente in A bezeichnen, und hat die Form

$$g_k = A x_k + B_{k\,r+1}\, x_{r+1} + B_{k\,r+2} x_{r+2} + \cdots + B_{kn}\, x_n + B_{k0},$$

wobei die Koeffizienten allgemein als Determinanten charakterisirt sind, nämlich es ist:

$$B_{kh} = \begin{vmatrix} a_{11} \dots a_{1\,k-1} & a_{1h} & a_{1\,k+1} \dots a_{1r} \\ a_{21} \dots a_{2\,k-1} & a_{2h} & a_{2\,k+1} \dots a_{2r} \\ \vdots & & \\ a_{r1} \dots a_{r\,k-1} & a_{rh} & a_{r\,k+1} \dots a_{rr} \end{vmatrix} = \sum_g A_{kg}\, a_{gh}.$$

$$(k, g = 1, 2, \dots r;\ h = 0, r+1 \dots n)$$

Das so definierte Modulsystem

$$(g_1, g_2, \dots g_r)$$

hat r von einander unabhängige Elemente und überhaupt die Form, auf die wir früher das lineare Modulsystem reduziert haben. Seine Elemente gehen durch lineare Transformation aus $f_1, f_2, \dots f_r$ hervor, und um nun noch zu untersuchen, ob es wirklich als reduziertes Modulsystem von $(f_1, f_2, \dots f_m)$ betrachtet werden kann, bleibt nur noch übrig zu zeigen, wie $f_1, f_2, \dots f_m$ linear durch $g_1, g_2, \dots g_r$ ausgedrückt werden können. Das kann aber folgendermaſsen geschehen. Bilden wir nämlich

$$\sum_k a_{ik} g_k = A \sum_k a_{ik} x_k + \sum_{k,h} B_{kh}\, a_{ik}\, x_h + \sum_k a_{ik} B_{k0},$$

$$(i = 1, 2, \dots m;\ k = 1, 2 \dots r;\ h = r+1, \dots n)$$

und beachten wir, daſs

$$\sum_k B_{kh}\, a_{ik} = \sum_{k,g} A_{kg}\, a_{gh}\, a_{ik} = - \begin{vmatrix} a_{11} \dots a_{1r} & a_{1h} \\ a_{21} \dots a_{2r} & a_{2h} \\ \vdots & \\ a_{r1} \dots a_{rr} & a_{rh} \\ a_{i1} \dots a_{ir} & 0 \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} \dots a_{1r} & 0 \\ a_{21} \dots a_{2r} & 0 \\ \vdots & \\ a_{r1} \dots a_{rr} & 0 \\ 0 \ \dots 0 & a_{ih} \end{vmatrix} - \begin{vmatrix} a_{11} \dots a_{1r} & a_{1h} \\ a_{21} \dots a_{2r} & a_{2h} \\ \vdots & \\ a_{r1} \dots a_{rr} & a_{rh} \\ a_{i1} \dots a_{ir} & a_{ih} \end{vmatrix}$$

ist (§ 50 III), und dafs die zuletzt auftretende Determinante auch für $h = r + 1, r + 2, \dots n$ verschwindet, weil der Rang des verkürzten Systems (a_{ik}) gleich r ist, so ergiebt sich die Gleichung

$$\sum_k a_{ik} g_k = A\left(\sum_h a_{ih} x_h + a_{i0}\right) - \begin{vmatrix} a_{11} \dots a_{1r} & a_{10} \\ \vdots & \\ a_{r1} \dots a_{rr} & a_{r0} \\ a_{i1} \dots a_{ir} & a_{i0} \end{vmatrix}$$

$$= A f_i - \begin{vmatrix} a_{11} \dots a_{1r} & a_{10} \\ \vdots & \\ a_{r1} \dots a_{rr} & a_{r0} \\ a_{i1} \dots a_{ir} & a_{i0} \end{vmatrix}, \quad \begin{pmatrix} i = 1, 2, \dots m; \\ k = 1, 2, \dots r; \\ h = 1, 2, \dots n \end{pmatrix}$$

die $A f_i$ und, da $A \neq 0$ ist, auch f_i als lineare Form der Gröfsen g_k und der Determinanten

$$\begin{vmatrix} a_{11} \dots a_{1r} & a_{10} \\ a_{21} \dots a_{2r} & a_{20} \\ \vdots & \\ a_{r1} \dots a_{rr} & a_{r0} \\ a_{i1} \dots a_{ir} & a_{i0} \end{vmatrix} \qquad (i = r + 1, \dots n)$$

darstellen lehrt. Sind diese letzteren alle gleich 0, so ist $(g_1, g_2, \dots g_r)$ in Wahrheit das reduzierte Modulsystem, ist aber auch nur eine von ihnen von 0 verschieden, so mufs dem System $(g_1 \dots g_r)$ eine von 0 verschiedene Konstante a beigefügt werden, damit es als ein reduziertes Modulsystem betrachtet werden kann.

Der Umstand, dafs der Rang des verkürzten Matrix (a_{ik}) gleich r ist, ist bei der vorstehenden Untersuchung nur dadurch zum Ausdruck gekommen, dafs aufser der Bedingung $A \neq 0$ nur die Determinanten

$$\begin{vmatrix} a_{11} & a_{12} \dots a_{1r} & a_{1k} \\ a_{21} & a_{22} \dots a_{2r} & a_{2k} \\ \vdots & & \\ a_{r1} & a_{r2} \dots a_{rr} & a_{rk} \\ a_{i1} & a_{i2} \dots a_{ir} & a_{ik} \end{vmatrix} = 0 \quad \begin{pmatrix} i = r + 1, \dots m; \\ k = r + 1 \dots n \end{pmatrix}$$

gesetzt wurden. Es ergiebt sich hieraus, dafs das Verschwinden dieser $(m - r)(n - r)$ Determinanten $(r + 1)$ Ordnung auch das Verschwinden aller übrigen — im ganzen giebt es $m_{r+1} n_{r+1}$ — nach sich zieht. Diese Determinanten sind aber wirklich unabhängig von einander, da in allen verschiedene Gröfsen a_{ik} vorhanden sind.

§ 53. Auflösung linearer Gleichungssysteme. Interpolationsformel von Lagrange.

I) Wie wir schon in § 46 gesehen haben, führt die Reduktion der linearen Modulsysteme unmittelbar zur Auflösung der linearen Gleichungssysteme. Auf Grund der Betrachtungen des vorigen Paragraphen sind wir nun imstande, die Lösungen des Gleichungssystems

$$f_1 = 0,\ f_2 = 0, \ldots f_m = 0$$

in der Form

$$g_1 = 0,\ g_2 = 0,\ g_r = 0$$

hinzuschreiben, die das ganze Problem auf die Berechnung von Determinanten zurückführt, vorausgesetzt, dafs der Rang der verkürzten Matrix mit dem der vollständigen übereinstimmt, denn sonst ist das Gleichungssystem nicht lösbar (§ 46). Wir erhalten allgemein

$$x_k = -\frac{B_{k\,r+1}\,x_{r+1} + B_{k\,r+2}\,x_{r+2} + \ldots + B_{kn}\,x_n + B_{k0}}{A}, \qquad (k = 1, 2, \ldots r)$$

wo $x_{r+1}, x_{r+2}, \ldots x_n$ ganz beliebige Werte annehmen können. Einige besondere Fälle wollen wir nun noch kurz erwähnen.

1) Wenn ein System von n linearen homogenen Gleichungen mit n Unbekannten eine nicht verschwindende Determinante hat, so haben sämtliche Unbekannte den Wert Null.

Giebt es also von Null verschiedene Lösungen, so mufs die Determinante des Gleichungssystems verschwinden.

2) Wenn ein System von n linearen Gleichungen mit n Unbekannten

$$\sum_k a_{ik}\,x_k = a_i \qquad (i, k = 1, 2, \ldots n)$$

eine von Null verschiedene Determinante besitzt, so sind die Unbekannten eindeutig bestimmt durch

$$x_k = \frac{\begin{vmatrix} a_{11} \ldots a_{1\,k-1} & a_1 & a_{1\,k+1} \ldots a_{1n} \\ \vdots & & \\ a_{n1} \ldots a_{n\,k-1} & a_n & a_{n\,k+1} \ldots a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} \ldots a_{1n} \\ \vdots \\ a_{n1} \ldots a_{nn} \end{vmatrix}}. \qquad (k = 1, 2, \ldots n)$$

Ist dagegen die Determinante des Gleichungssystems gleich Null, so müssen auch die im Zähler auftretenden n Determinanten sicher verschwinden, wenn das Gleichungssystem lösbar sein soll.

II. Wir wollen nun die dargelegten Entwickelungen dazu benutzen, um zu untersuchen, durch wieviel Werte eine ganze rationale Funktion $n$ten Grades

$$F(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n$$

bestimmt ist. Sind $x_0, x_1, x_2, \dots x_m$ eine Reihe von verschiedenen Werten, so gelten die Gleichungen

$$F(x_i) = a_0 x_i^n + a_1 x_i^{n-1} + \dots + a_{n-1} x_i + a_n, \qquad (i = 0, 1, 2, \dots m)$$

die wir als ein System linearer Gleichungen in den $(n + 1)$ Gröfsen

$$a_0, a_1, \dots a_n$$

betrachten können. Damit diese völlig bestimmt sind, mufs m mindestens gleich $n + 1$ sein und die Determinante

$$|x_i^k| \neq 0 \qquad (i, k = 0, 1, \dots n)$$

sein. Nun läfst sich aber diese Determinante (§ 54 Beisp. 24) als Produkt aller Differenzen der Gröfsen $x_0, x_1 \dots x_n$ darstellen und ist daher sicher von Null verschieden nach unserer Annahme. Also ist eine ganze Funktion $n$ten Grades völlig bestimmt, wenn die Werte gegeben sind, die sie für $(n + 1)$ beliebig gewählte verschiedene Werte von x annimmt. Die Werte der Koeffizienten lassen sich leicht darstellen, aber es mag dem Leser überlassen bleiben, dies durchzuführen. Da es uns aber darauf ankommt, $F(x)$ selbst herzustellen, so ziehen wir es vor, das System

$$\begin{aligned} F(x) &= a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n \\ F(x_i) &= a_0 x_i^n + a_1 x_i^{n-1} + \dots + a_{n-1} x_i + a_n \end{aligned} \qquad (i = 0, 1, \dots n)$$

als ein in den nicht sämtlich verschwindenden Gröfsen

$$1, a_0, a_1, \dots a_{n-1}, a_n$$

homogenes lineares Gleichungssystem aufzufassen, dessen Determinante daher gleich 0 sein mufs. Aus der Gleichung

$$\begin{vmatrix} F(x) & x^n & x^{n-1} & \dots & x & 1 \\ F(x_0) & x_0^n & x_0^{n-1} & \dots & x_0 & 1 \\ F(x_1) & x_1^n & x_1^{n-1} & \dots & x_1 & 1 \\ \vdots \\ F(x_n) & x_n^n & x_n^{n-1} & \dots & x_n & 1 \end{vmatrix} = 0$$

folgt nun aber durch Entwickelung der Determinante nach der ersten Spalte, wenn man noch zur Abkürzung

$$M(x) = (x - x_0)(x - x_1) \dots (x - x_n)$$

einführt und das Beispiel 24 in § 51 beachtet,

$$\frac{F(x)}{M(x)} - \sum_{k=0}^{k=n} \frac{F(x_k)}{(x - x_k) M'(x_k)} = 0$$

oder

$$F(x) = \sum_{k=0}^{k=n} F(x_k) \frac{M(x)}{(x - x_k) M'(x_k)}.$$

Dieses ist die Interpolationsformel von Lagrange, die eine Funktion n ten Grades konstruieren lehrt, die für $n + 1$ beliebige verschiedene Werte $x_0, x_1 \dots x_n$ von x selbst die Werte $F(x_0), F(x_1), \dots F(x_n)$ annimmt.

---

§ 54. Multiplikationstheorem der Determinanten.

Ersetzt man in einer Determinante alle Elemente einer Zeile oder Spalte durch dieselbe lineare Form bezw. der Elemente der Spalten oder Zeilen, so läfst sich die Änderung der Determinante leicht beurteilen. Man hat z. B. die Gleichung

$$\begin{vmatrix} a_{11} b_1 + a_{12} b_2 + \dots + a_{1n} b_n & a_{12} \dots a_{1n} \\ a_{21} b_1 + a_{22} b_2 + \dots + a_{2n} b_n & a_{22} \dots a_{2n} \\ \vdots \\ a_{n1} b_1 + a_{n2} b_2 + \dots + a_{nn} b_n & a_{n2} \dots a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \vdots \\ a_{n1} & a_{n2} \dots a_{nn} \end{vmatrix} b_1.$$

Setzt man

$$\begin{aligned} f &= b_1 x_1 + b_2 x_2 + \dots + b_n x_n \\ f' &= b'_1 x_1 + b'_2 x_2 + \dots + b'_n x_n, \end{aligned}$$

so ergiebt sich wieder

$$\begin{vmatrix} f(a_{11}, a_{12}, \dots a_{1n}) & f'(a_{11}, a_{12}, \dots a_{1n}) & a_{13} \dots a_{1n} \\ f(a_{21}, a_{22}, \dots a_{2n}) & f'(a_{21}, a_{22}, \dots a_{2n}) & a_{23} \dots a_{2n} \\ \vdots & & \\ f(a_{n1}, a_{n2}, \dots a_{nn}) & f'(a_{n1}, a_{n2}, \dots a_{nn}) & a_{n3} \dots a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} b_1 + a_{12} b_2 & a_{11} b'_1 + a_{12} b'_2 & a_{13} \dots a_{1n} \\ a_{21} b_1 + a_{22} b_2 & a_{21} b'_1 + a_{22} b'_2 & a_{23} \dots a_{2n} \\ \vdots & & \\ a_{n1} b_1 + a_{n2} b_2 & a_{n1} b'_1 + a_{n2} b'_2 & a_{n3} \dots a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} \dots a_{1n} \\ \vdots \\ a_{n1} \dots a_{nn} \end{vmatrix} (b_1 b'_2 - b_2 b'_1).$$

In dieser Weise kann man fortfahren, und die allgemeinste Art der Umformung ist offenbar die, daſs jede Reihe durch lineare Formen ersetzt wird. Setzen wir demgemäſs

$$f_i = b_{1i} x_1 + b_{2i} x_i + \dots + b_{ni} x_n$$

und betrachten nun die Determinante

$$\begin{vmatrix} f_1(a_{11}, \dots a_{1n}) & f_2(a_{11}, \dots a_{1n}) \dots & f_n(a_{11}, \dots a_{1n}) \\ f_1(a_{21}, \dots a_{2n}) & f_2(a_{21}, \dots a_{2n}) & f_n(a_{2n}, \dots a_{2n}) \\ \vdots & & \vdots \\ f_1(a_{n1}, \dots a_{nn}) & f_2(a_{n1}, \dots a_{nn}) & f_n(a_{n1}, \dots a_{nn}) \end{vmatrix} = | c_{ik} |,$$

so läſst sich diese auch offenbar so charakterisieren, daſs ihre Matrix (c_{ik}) aus der Kompostion der beiden Matrizen

$$(a_{ik}) \text{ und } (b_{ik}) \qquad (i, k = 1, 2, \dots n)$$

hervorgeht. Durch successive Umformung erhält man aus der Determinante

$$\sum_{i_1} \begin{vmatrix} a_{1i_1} & f_2(a_{11}, \dots a_{1n}) \dots f_n(a_{11}, \dots a_{1n}) \\ a_{2i_1} & f_2(a_{21}, \dots a_{2n}) \dots f_n(a_{21}, \dots a_{2n}) \\ \vdots & \\ a_{ni_1} & f_n(a_{n1}, \dots a_{nn}) \dots f_n(a_{n1}, \dots a_{nn}) \end{vmatrix} b_{i_1 1}$$

$$= \sum_{i_1, i_2} \begin{vmatrix} a_{1i_1} & a_{1i_2} & f_3(a_{11}, \dots a_{1n}) \dots f_n(a_{11}, \dots a_{1n}) \\ a_{2i_1} & a_{2i_2} & f_3(a_{21}, \dots a_{2n}) \dots f_n(a_{21}, \dots a_{2n}) \\ \vdots & & \\ a_{ni_1} & a_{ni_2} & f_n(a_{n1}, \dots a_{nn}) \dots f_n(a_{n1}, \dots a_{nn}) \end{vmatrix} b_{i_1 1} b_{i_2 2}.$$

Schliesslich ergiebt sich

$$\sum_{i_1, i_2 \ldots i_n} \begin{vmatrix} a_{1 i_1} & a_{1 i_2} \ldots a_{1 i_n} \\ a_{2 i_1} & a_{2 i_2} \ldots a_{2 i_n} \\ \vdots & \\ a_{n i_1} & a_{n i_2} \ldots a_{n i_n} \end{vmatrix} b_{i_1 1} b_{i_2 2} \ldots b_{i_n n}.$$

Nun ist aber

$$\begin{vmatrix} a_{1 i_1} & a_{1 i_2} \ldots a_{1 i_n} \\ a_{2 i_1} & a_{2 i_2} \ldots a_{2 i_n} \\ \vdots & \\ a_{n i_1} & a_{n i_2} \ldots a_{n i_n} \end{vmatrix} = \varepsilon_{i_1 i_2 \ldots i_n} \begin{vmatrix} a_{11} \ldots a_{1n} \\ a_{21} \ldots a_{2n} \\ \vdots \\ a_{n1} \ldots a_{nn} \end{vmatrix},$$

$$\sum_{i_1, i_2 \ldots i_n} \varepsilon_{i_1 i_2 \ldots i_n} b_{i_1 1} b_{i_2 2} \ldots b_{i_n n} = \begin{vmatrix} b_{11} \ldots b_{1n} \\ \vdots \\ b_{n1} \ldots b_{nn} \end{vmatrix},$$

so dafs unsere Determinante einfach gleich

$$|a_{ik}| \,.\, |b_{ik}| \qquad (i, k = 1, 2, \ldots n)$$

ist. Wir erhalten so das sog. Multiplikationstheorem der Determinante in der Form:

Die Determinante des komponierten Systems ist gleich dem Produkte der Determinanten der Komponenten.

Nicht wesentlich verschieden von der soeben dargelegten Ableitung dieses Satzes ist die folgende, bei der von vornherein von der Summendefinition der Determinante Gebrauch gemacht wird.

$$|c_{ik}| = \sum_{h_1, h_2 \ldots h_n} \varepsilon_{h_1 h_2 \ldots h_n} c_{1 h_1} c_{2 h_2} \ldots c_{n h_n}$$

$$= \sum_{h_1, h_2 \ldots h_n, i_1, i_2 \ldots i_n} \varepsilon_{h_1 h_2 \ldots h_n} a_{1 i_1} b_{i_1 h_1} a_{2 i_2} b_{i_2 h_2} \ldots a_{n i_n} b_{i_n h_n}$$

$$= \sum_{h_1, \ldots h_n,\ i_1, i_2 \ldots i_n} \varepsilon_{h_1 \ldots h_n} b_{i_1 h_1} b_{i_2 h_2} \ldots b_{i_n h_n} \,.\, a_{1 i_1} a_{2 i_2} \ldots a_{n i_n}$$

$$= \sum_{i_1, i_2 \ldots i_n} \begin{vmatrix} b_{i_1 1} & b_{i_1 2} \ldots b_{i_1 n} \\ b_{i_2 1} & b_{i_2 2} \ldots b_{i_2 n} \\ \vdots & \\ b_{i_n 1} & b_{i_n 2} \ldots b_{i_n n} \end{vmatrix} a_{1 i_1} a_{2 i_2} \ldots a_{n i_n}$$

$$= \sum_{i_1, i_2 \ldots i_n} \varepsilon_{i_1 i_2 \ldots i_n} a_{1 i_1} a_{2 i_2} \ldots a_{n i_n} \begin{vmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \\ b_{n1} & \ldots & b_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{n1} & \ldots & a_{nn} \end{vmatrix} \cdot \begin{vmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & & \\ b_{n1} & \ldots & b_{nn} \end{vmatrix}.$$

---

§ 55. Anwendungen des Multiplikationstheorems.

I. Die Komposition der Matrizen bot sich uns ursprünglich bei der Zusammensetzung linearer Transformationen in § 10 dar; dann haben wir in § 11 gezeigt, dafs sie auch bei der linearen Transformation quadratischer Formen auftritt, und endlich fanden wir in § 50 die Komposition jeder Matrix mit der adjungierten Matrix. Auf alle diese Fälle können wir das Multiplikationstheorem der Determinanten in Anwendung bringen und gelangen so zu einer Reihe von Sätzen, die wir zunächst betrachten wollen.

1) Wenn ein System von n Linearformen

$$f_i = \sum_k a_{ik} x_k \qquad (i, k = 1, 2, \ldots n)$$

durch die lineare Transformation

$$x_i = \sum_k \alpha_{ik} x_k' \qquad (i, k = 1, 2, \ldots n)$$

in das System

$$f'_i = \sum_k a'_{ik} x'_k \qquad (i, k = 1, 2, \ldots n)$$

transformiert wird, so ist

$$|a'_{ik}| = |a_{ik}| \cdot |\alpha_{ik}|. \qquad (i, k = 1, 2, \ldots n)$$

Von diesem Satze haben wir schon früher bei Behandlung der Kettenbrüche einen einfachen Fall kennen gelernt. Vgl. § 13.

2) Die Determinante einer durch lineare Transformation der Variabeln entstehenden quadratischen Form ist gleich dem Produkt aus der Determinante der ursprünglichen Form und dem Quadrat der Determinante der linearen Transformation.

3) Die Determinante des adjungierten Systems von n^2 Elementen ist gleich der $(n-1)$ten Potenz der Determinante dieser Elemente.

Aus § 50 I ergiebt sich nämlich sofort

$$|A_{ik}| \cdot |a_{ik}| = |a_{ik}|^n, \qquad (i, k = 1, 2 \ldots n)$$

hieraus aber durch die Division mit $|a_{ik}|$

$$|A_{ik}| = |a_{ik}|^{n-1}, \qquad (i, k = 1, 2 \ldots n)$$

wobei allerdings zunächst vorausgesetzt wird, dafs $|a_{ik}| \neq 0$ ist. Da aber die zuletzt entstehende Gleichung eine identische sein mufs, so gilt sie auch dann noch, wenn $|a_{ik}| = 0$ ist.

II. Das Multiplikationstheorem lehrt nur die Darstellung eines Produktes zweier Determinanten gleicher Ordnung kennen. Sind aber zwei Determinanten verschiedener Ordnung vorgelegt, so kann man die Determinante von kleinerer Ordnung leicht in eine höherer Ordnung verwandeln, indem man z. B. die Diagonale durch Einfügung lauter Einheiten verlängert, auf der einen Seite dieser Diagonale lauter Nullen, auf der andern aber beliebige Gröfsen hinsetzt. Vgl. § 51 Beispiele 10 und 11. Ein Beispiel fügen wir zur Veranschaulichung bei.

$$4) \quad \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \begin{vmatrix} e_1 & f_1 \\ e_1 & f_2 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \begin{vmatrix} e_1 & f_1 & 0 & 0 \\ e_2 & f_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 e_1 + b_1 e_2 & a_1 f_1 + b_1 f_2 & c_1 & d_1 \\ a_2 e_1 + b_2 e_2 & a_2 f_1 + b_2 f_2 & c_2 & d_2 \\ a_3 e_1 + b_3 e_2 & a_3 f_1 + b_3 f_2 & c_3 & d_3 \\ a_4 e_1 + b_4 e_4 & a_4 f_1 + b_4 f_2 & e_4 & d_4 \end{vmatrix}.$$

III. Wir wollen nun noch einige Beispiele für das Multiplikationstheorem beifügen.

$$5) \qquad \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}^2$$

$$= \begin{vmatrix} a_1^2 + b_1^2 + c_1^2 & a_1 a_2 + b_1 b_2 + c_1 c_2 & a_1 a_3 + b_1 b_3 + c_1 c_3 \\ a_1 a_2 + b_1 b_2 + c_1 c_2 & a_2^2 + b_2^2 + c_2^2 & a_2 a_3 + b_2 b_3 + c_2 c_3 \\ a_1 a_3 + b_1 b_3 + c_1 c_3 & a_2 a_3 + b_2 b_3 + c_2 c_3 & a_3^2 + b_3^2 + c_3^2 \end{vmatrix}.$$

6) Um das Quadrat der in § 51 als Beispiel 24 behandelten Determinante

$$|x_i^{k-1}| \qquad (i, k = 1, 2, \ldots n)$$

in einfacher Weise zu berechnen, komponieren wir die Matrix (x_i^{k-1}) mit der transponierten Matrix (§ 11).

Setzen wir

$$s_0 = n, \qquad s_i = x_1^i + x_2^i + \ldots + x_n^i,$$

so ergiebt sich

$$\begin{pmatrix} 1 & 1 & \ldots 1 \\ x_1 & x_2 & \ldots x_n \\ \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots x_n^{n-1} \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^2 \ldots x_1^{n-1} \\ 1 & x_2 & x_2^2 \ldots x_2^{n-1} \\ \vdots \\ 1 & x_n & x_n^2 \ldots x_n^{n-1} \end{pmatrix} = \begin{pmatrix} s_0 & s_1 \ldots s_{n-1} \\ s_1 & s_2 \ldots s_n \\ \vdots \\ s_{n-1} & s_n \ldots s_{2n-2} \end{pmatrix}.$$

Daher ist

$$| x_i^{k-1} |^2 = | s_{i+k-2} |. \qquad (i, k = 1, 2, \ldots n)$$

Weitere Beispiele für die Multiplikation der Determinanten sind in dem neunten Abschnitt über Resultanten enthalten, auf den wir hiermit verweisen.

VII. Abschnitt.

Teilbarkeit ganzer Funktionen.

§ 56. Division ganzer Funktionen.

Wir nehmen an, dafs zwei ganze Funktionen A(x) und B(x) einer Veränderlichen x von den Graden α und β mit Koeffizienten eines beliebigen Rationalitätsbereiches (§ 4) vorliegen. Es sei $\alpha > \beta$, und die nicht verschwindenden Koeffizienten der in ihnen vorkommenden höchsten Potenzen von x mögen a_0 und b_0 heifsen. Es ist dann leicht, aus A(x) durch Subtraktion eines Produktes aus B(x) eine Funktion $A_1(x)$ herzustellen, deren Grad $\alpha_1 < \alpha$ ist. Setzt man nämlich $\frac{a_0}{b_0} = g_0$ und bildet

$$A(x) - g_0 x^{\alpha - \beta} B(x) = A_1(x),$$

so hebt sich das Glied mit der Potenz x^α weg, und es entsteht also eine Funktion, deren Grad α_1 höchstens gleich $\alpha - 1$ sein kann. Ist $\alpha_1 > \beta$, so kann das Verfahren fortgesetzt werden und so oft, bis eine neue Funktion zum Vorschein kommt, deren Grad kleiner als β ist, was nach einer beschränkten Anzahl von Bildungen der beschriebenen Art sicher eintritt, da die Grade der gebildeten Funktionen beständig abnehmen. Man erhält so eine Kette von Gleichungen, von denen die erste oben hingeschrieben ist, die folgenden die Form haben:

$$\begin{aligned} A_1(x) &- g_1 x^{\alpha_1 - \beta} B(x) = A_2(x) \\ A_2(x) &- g_2 x^{\alpha_2 - \beta} B(x) = A_3(x) \\ &\vdots \\ A_{i-1}(x) &- g_{i-1}^{\alpha_{i-1} - \beta} B(x) = A_i(x), \end{aligned}$$

und die letzte, mit der das Verfahren seinen Abschlufs findet, in der Gestalt

$$A_i(x) - g_i x^{\alpha_i - \beta} B(x) = C(x)$$

geschrieben werden soll, wobei der Grad γ von $C(x)$ kleiner als β ist. Aus allen diesen Gleichungen ergiebt sich, wenn man die ganze Funktion

$$G(x) = g_0 x^{\alpha - \beta} + g_1 x^{\alpha_1 - \beta} + \ldots + g_i x^{\alpha_i - \beta}$$

einführt, durch Addition, dafs

$$A(x) - G(x) B(x) = C(x)$$

ist, und damit der Satz:

Zu zwei ganzen Funktionen $A(x)$ und $B(x)$ kann man stets und nur auf eine einzige Weise zwei neue ganze Funktionen $G(x)$ und $C(x)$ bestimmen, dafs

$$A(x) = G(x) B(x) + C(x)$$

und der Grad von $C(x)$ kleiner als der von $B(x)$ ist, wobei die Koeffizienten der neuen Funktionen rational mit denen von $A(x)$ und $B(x)$ zusammenhängen.

Man nennt das dargelegte Verfahren auch oft Division, $G(x)$ den Quotienten, $C(x)$ den Rest der Division von $A(x)$ durch $B(x)$. Schreibt man die zuletzt erhaltene Gleichung in der Form

$$\frac{A(x)}{B(x)} = G(x) + \frac{C(x)}{B(x)},$$

so kann man ihren Inhalt auch so aussprechen: Jede gebrochene rationale Funktion einer Variabeln kann immer und nur auf eine einzige Weise als Summe einer ganzen Funktion und solchen gebrochenen Funktion dargestellt werden, dessen Zähler einen niederen Grad besitzt als der Nenner. Rationale Funktionen der letzteren Art kann man als echt gebrochene bezeichnen. Ist die Funktion $\frac{A(x)}{B(x)}$ schon echt gebrochen, so verschwindet $G(x)$ und ist $A(x) = C(x)$.

Ergiebt sich bei dem dargelegten Verfahren ein verschwindender Rest, existiert also zu A(x) nnd B(x) eine ganze Funktion C(x), so dafs

$$A(x) = B(x)\,C(x)$$

ist, so sagt man von A(x), es sei durch B(x) teilbar, von B(x), dafs es ein Teiler von A(x) sei oder in A(x) aufgehe. Wenn A(x) durch B(x) teilbar ist, so ist die dann eindeutig bestimmte Funktion $C(x) = A(x) : B(x)$ ebenfalls ein Teiler von A(x) und wird als der zu B(x) komplementäre Teiler von A(x) bezeichnet. Sehen wir von ganzen Funktionen mit verschwindenden Koeffizienten ab, die offenbar durch jede beliebige Funktion teilbar sind, so ist der Grad des Teilers stets kleiner oder höchstens gleich dem Grad der Funktion, in der er enthalten ist. Tritt das Letztere ein, so ist der komplementäre Teiler eine Konstante d. h. eine von x unabhängige Gröfse, und eine solche Konstante ist, wenn sie nicht verschwindet, immer Teiler jeder beliebigen ganzen Funktion. Daher hat eine ganze Funktion unendlich viele Teiler, und man kann überdies auch noch jeden Teiler, der eine wirkliche Funktion von x ist, mit einer beliebigen nicht verschwindenden Konstanten multiplizieren, ohne dafs er seine Teilereigenschaft dadurch verliert. Aber wenn man übereinkommt, ganze Funktionen, die sich nur um eine multiplikative Konstante von einander unterscheiden, nicht als verschieden anzusehen, so läfst sich zeigen, dafs jede ganze Funktion nur eine beschränkte Anzahl von Teilern besitzen kann. Um diesen Beweis zu ermöglichen, vorläufig für ganzzahlige Funktionen, sind die folgenden Betrachtungen nötig, die auch an sich von grofser Wichtigkeit sind.

---

§ 57. Produkte ganzer Funktionen.

Summen und Differenzen von ganzen Funktionen sind wieder ganze Funktionen, ebenso auch Produkte, und zwar ist hierbei der Grad des Produktes gleich der Summe der Grade der Faktoren. Sind A(x) und B(x) zwei ganze Funktionen:

$$A(x) = a_0 x^\alpha + a_1 x^{\alpha-1} + \ldots + a_{\alpha-1} x + a_\alpha$$
$$B(x) = b_0 x^\beta + b_1 x^{\beta-1} + \ldots + b_{\beta-1} x + b_\beta$$

von den Graden α und β, so ist ihr Produkt

$$C(x) = c_0 x^\gamma + c_1 x^{\gamma-1} + \ldots + c_{\gamma-1} x + c_\gamma$$

von Grade $\gamma = \alpha + \beta$, wo die Koeffizienten durch die folgenden Gleichungen bestimmt sind

$$\begin{aligned} c_0 &= a_0 b_0 \\ c_1 &= a_0 b_1 + a_1 b_0 \\ &\vdots \\ c_\gamma &= a_\alpha b_\beta, \end{aligned}$$

und allgemein ist für $h = 0, 1 \ldots \gamma$

$$c_h = \Sigma a_i b_k, \qquad (i = 0, 1 \ldots \alpha;\ k = 0, 1, \ldots \beta)$$

wenn die Summe rechts über alle Koeffizienten a_i, b_k erstreckt wird, deren Indizes der Bedingung $h = i + k$ genügen.

Wir nehmen nun an, dafs die Koeffizienten von $A(x)$ und $B(x)$ ganze Zahlen sind, dann ist auch $C(x)$ eine ganze ganzzahlige Funktion. Führen wir dann für die gröfsten gemeinschaftlichen Teiler der Koeffizienten die Bezeichnungen

$$\begin{aligned} a &= (a_0, a_1, \ldots a_\alpha) \\ b &= (b_0, b_1, \ldots b_\beta) \\ c &= (c_0, c_1, \ldots c_\gamma) \end{aligned}$$

ein, so läfst sich nachweisen, dafs

$$a\,b = c$$

ist. Zunächst ist unmittelbar aus der Darstellung von c_h als Summe von lauter durch $a\,b$ teilbaren Summanden ersichtlich, dafs alle Koeffizienten von $C(x)$ und daher auch $c = (c_0, c_1 \ldots c_\gamma)$ durch $a\,b$ teilbar sind. Es braucht daher nur noch nachgewiesen zu werden, dafs sie aufser $a\,b$ keinen gemeinschaftlichen Teiler besitzen, oder dafs es keine Primzahl p giebt, durch die alle Koeffizienten $\frac{c_h}{a\,b}$ teilbar sind. Da alle Quotienten $\frac{a_i}{a}$ und $\frac{b_i}{b}$ die Einheit als gröfsten gemeinschaftlichen Teiler haben, so kann eine solche Primzahl p nicht in allen von ihnen, sondern nur in einigen als Teiler

enthalten sein. Nehmen wir daher an, es sei p enthalten in $\frac{a_0}{a}, \frac{a_1}{a}, \ldots \frac{a_{i-1}}{a}$, aber nicht in $\frac{a_i}{a}$, ferner in $\frac{b_0}{b}, \frac{b_1}{b}, \ldots \frac{b_{k-1}}{b}$, aber nicht in $\frac{b_k}{b}$, so ergiebt sich aus:

$$\frac{c_{i+k}}{ab} = \frac{a_i b_k}{ab} + \frac{a_{i-1}}{a} \frac{b_{k+1}}{b} + \frac{a_{i-2}}{a} \frac{b_{k+2}}{b} + \ldots$$
$$+ \frac{a_{i+1}}{a} \frac{b_{k-1}}{b} + \frac{a_{i+2}}{a} \frac{b_{k-2}}{b} + \ldots,$$

da alle Glieder der rechten Seite mit alleiniger Ausnahme von $\frac{a_i}{a} \frac{b_k}{b}$ durch p teilbar sind, dafs $\frac{c_{i+k}}{ab}$ nicht durch p teilbar sein kann. Da also alle Koeffizienten c_h aufser ab nicht noch irgend eine Primzahl als Teiler besitzen können, so ist ab der gröfste gemeinschaftliche Teiler aller Koeffizienten c_h.

Nennt man den gröfsten gemeinschaftlichen Teiler der ganzzahligen Koeffizienten einer ganzen Funktion kurz deren Zahlenteiler, so kann die soeben bewiesene Thatsache auch so formuliert werden:

Der Zahlenteiler des Produktes von ganzen ganzzahligen Funktionen ist gleich dem Produkte der Zahlenteiler der einzelnen Faktoren.

Haben die Koeffizienten einer ganzen ganzzahligen Funktion keinen andern gemeinschaftlichen Teiler als die Einheit, so nennt man die Funktion primitiv. Also ist das Produkt von primitiven Funktionen wieder eine primitive Funktion.

Wichtige Folgerungen lassen sich aus diesen Sätzen über die Zerlegung ganzer ganzzahliger Funktionen ziehen. Nehmen wir an, dafs C(x) ganzzahlig sei, aber als das Produkt $A_1(x)\,B_1(x)$ zweier zwar ganzer, aber nicht ganzzahliger Funktionen $A_1(x)$ und $B_1(x)$ darstellbar sei, so kann man

$$A_1(x) = \frac{A'(x)}{n}, \quad B_1(x) = \frac{B'(x)}{m}$$

setzen, wo A'(x) und B'(x) nun auch ganzzahlige Funktionen sind; n und m kann man dabei als kleinste gemeinschaft-

liche Vielfache der in den Koeffizienten auftretenden Nenner (Generalnenner) ansehen. Aus der Gleichung

$$A'(x)\,B'(x) = n\,m\,C(x)$$

folgt dann aber, wenn a′, b′, c die Zahlenteiler von A′(x), B′(x), C(x) bedeuten, vermöge des bewiesenen Satzes, dafs

$$a'b' = n\,m\,c$$

ist. Zerlegt man nun c in zwei Faktoren

$$c = e\,f$$

und setzt

$$A(x) = e\,n\,\frac{A'(x)}{a'}, \quad B(x) = f\,m\,\frac{B'(x)}{b'},$$

so sind A(x) und B(x) ganze ganzzahlige Funktionen, die der Gleichung

$$A(x)\,B(x) = C(x)$$

genügen. Läfst sich also eine ganzzahlige Funktion in ein Produkt von ganzen Funktionen mit rationalen Koeffizienten darstellen, so kann man sie auch als ein Produkt von ganzzahligen Funktionen zerlegen, die sich von ihnen nur durch multiplikative Konstanten unterscheiden. Wenn also eine ganze ganzzahlige Funktion einen Teiler besitzt, so darf nicht nur dieser, sondern auch der komplementäre Teiler als ganzzahlig vorausgesetzt werden.

Besonders bemerkenswert ist nun noch der Fall, wo der Koeffizient der höchsten Potenz in C(x) die Einheit ist. Alsdann können zwei komplementäre Teiler A(x) und B(x) offenbar dadurch völlig bestimmt werden, dafs in ihnen die Koeffizienten der höchsten Potenzen auch gleich der Einheit sein sollen. Es ergeben sich folgende Sätze: Ist eine ganze ganzzahlige Funktion durch eine ganze Funktion teilbar, und besitzt sie als Koeffizienten der höchsten Potenz die Einheit, so ist der Teiler auch ganzzahlig, ebenso wie der komplementäre. Sind in zwei ganzen Funktionen die Koeffizienten der höchsten Potenz Einheiten, so kann ihr Produkt nicht ganzahlig sein, wenn die beiden Funktionen nicht auch ganzzahlig waren.

§ 58. Anzahl der Teiler ganzer Funktionen. Primfunktionen.

Da jede ganze Funktion einer Veränderlichen mit rationalen Koeffizienten durch jede beliebige Konstante und sich selbst teilbar ist, und ferner jedem andern Teiler, wie wir soeben bewiesen haben, immer ein Teiler mit ganzzahligen Koeffizienten ohne einen andern wesentlichen Teiler, als die Einheit entspricht, so ist es zweckmäfsig, die folgenden Begriffe einzuführen. Wir nennen jede primitive ganze Funktion B(x), die ein Teiler von A(x) ist, aber deren Grad nicht erreicht, einen eigentlichen Teiler von A(x), und stellen ihr alle andern Teiler, deren es in unendlicher Anzahl giebt, als uneigentliche Teiler gegenüber. Dann läfst sich zeigen, dafs die Anzahl der eigentlichen Teiler einer ganzen Funktion mit rationalen Koeffizienten eine beschränkte ist. Offenbar genügt es, dies für die eigentlichen Teiler von einem bestimmten Grade zu zeigen, der kleiner als der der ganzen Funktion ist.

Die gegebene ganze Funktion F(x) vom Grade m können wir, wie leicht einzusehen ist, als ganzzahlig, sogar als primitiv voraussetzen, doch ist das Letztere für den Beweis nicht nötig. Wenn nun G(x) ein eigentlicher Teiler von F(x) ist, so ist der komplementäre Teiler als ganzzahlig anzusehen, und daraus folgt, dafs für jede ganze Zahl r die ganze Zahl F(r) auch durch G(r) teilbar sein mufs. Sei nun G(x) vom Grad $n < m$, so nehmen wir $(n + 1)$ verschiedene ganze Zahlen $r_0, r_1, \ldots r_n$ an, so dafs, wenn

$$G(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

ist, die $(n + 1)$ Gleichungen

$$G(r_i) = a_0 r_i^n + a_1 r_i^{n-1} + \ldots + a_{n-1} r_i + a_n$$
$$(i = 0, 1, \ldots n)$$

bestehen. Diese ermöglichen es nun, die Koeffizienten a_0, $a_1, \ldots a_n$ linear durch die ganzen Zahlen $G(r_i)$ auszudrücken in der Form:

$$a_k = h_{k0} G(r_0) + h_{k1} G(r_1) + \ldots + h_{kn} G(r_n),$$
$$(k = 0, 1 \ldots n)$$

wobei die $(n+1)^2$ Gröfsen h_{ik} nur von $r_0, r_1, \ldots r_n$ abhängig sind und daher gewisse von vornherein angebbare Werte nicht überschreiten können (vgl. § 53). Dasselbe gilt aber auch für die ganzen Zahlen $G(r_0), G(r_1), \ldots G(r_n)$, die beziehungsweise Teiler der festen ganzen Zahlen $F(r_0), F(r_1), \ldots F(r_n)$ sind. Daraus folgt unmittelbar, dafs auch die Koeffizienten a_k unterhalb gewisser Grenzen liegen müssen, und da sie ganze Zahlen sein sollen, nur in beschränkter Anzahl vorhanden sein können. Daher giebt es auch nur eine beschränkte Anzahl eigentlicher Teiler vom Grad n, und weil n nur die Werte $1, 2 \ldots m-1$ annehmen kann, auch nur eine beschränkte Anzahl von eigentlichen Teilern von $F(x)$ überhaupt.

Dieser Beweis zeigt zugleich einen Weg, auf dem man alle Teiler einer ganzen Funktion bestimmen könnte, obwohl derselbe sehr umständlich und in dieser Weise für die Praxis nicht von grofsem Wert ist. Was den Grad n des Teilers betrifft, so kann man sich wegen der Existenz der komplementären Teiler darauf beschränken $2n \leqq m$ anzunehmen.

Jede Funktion, die keine eigentlichen Teiler besitzt, nennt man eine Primfunktion oder eine irreduktible Funktion; alle andern (mit Ausschlufs der Konstanten) heifsen zusammengesetzte oder reduktible Funktionen. Jede zusammengesetzte Funktion hat also mindestens einen eigentlichen Teiler und läfst sich als Produkt von zwei Funktionen darstellen, die beide für sich eigentliche Teiler sind. Sobald eine von diesen wieder zusammengesetzt ist, kann man sie wieder zerlegen und so fortfahren, bis man auf Teiler stöfst, die nicht mehr eigentlich zerlegbar und daher Primfunktionen sind, was wegen der abnehmenden Gradzahlen oder auftretenden Funktionen sicher eintritt. Jede zusammengesetzte Funktion läfst sich also als ein Produkt von lauter Primfunktionen darstellen. Es geht aber aus dieser Ableitung nicht hervor, dafs diese Darstellung, abgesehen von der Reihenfolge der Faktoren, nur auf eine Art möglich ist. (Vgl. die analogen Betrachtungen in § 16.)

---

§ 59. Primfunktionen ersten Grades. Rationale Wurzeln einer algebraischen Gleichung.

Der niedrigste Grad, den eine Primfunktion haben kann, ist der erste, und umgekehrt sind alle Funktionen ersten Grades auch Primfunktionen. Sie haben die Form $(x - r)$, wo r eine rationale Zahl bedeutet. Wenn nun eine ganze Funktion $F(x)$ durch $(x - r)$ teilbar ist, so verschwindet sie, wenn man $x = r$ setzt, und auch umgekehrt, wenn $F(r) = 0$ ist, so ist $F(x)$ durch $(x - r)$ teilbar, da allgemein $F(x) - F(r)$ durch $(x - r)$ teilbar ist, wie wir jetzt zeigen wollen.

Um die Funktion $F(x)$, die in der Form

$$F(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

dargestellt sein möge, durch $(x - r)$ zu dividieren, kann man das in § 56 angegebene Verfahren einschlagen. Kürzer läſst sich das Resultat ermitteln, wenn wir den Quotienten

$$\frac{F(x) - F(r)}{x - r}$$

betrachten, der gleich

$$a_0 \frac{x^n - r^n}{x - r} + a_1 \frac{x^{n-1} - r^{n-1}}{x - r} + \ldots + a_{n-1}$$

ist. Beachten wir, daſs

$$\frac{x^i - r^i}{x - r} = x^{i-1} + x^{i-2} r + \ldots + x r^{i-2} + r^{i-1} \qquad (i = 1, 2 \ldots n)$$

ist, so ergiebt sich, wenn wir nach Potenzen von x ordnen, der Quotient

$$\frac{F(x) - F(r)}{x - r} = b_0 x^{n-1} + b_1 x^{n-2} + \ldots + b_{n-2} x + b_{n-1}$$

als eine ganze Funktion von x mit den Koeffizienten;

$$\begin{aligned} b_0 &= a_0 \\ b_1 &= a_0 r + a_1 \\ &\vdots \\ b_i &= a_0 r^i + a_1 r^{i-1} + \ldots a_{i-1} r + a_i \quad (i = 0, 1, \ldots n - 1) \\ &\vdots \\ b_{n-1} &= a_0 r^{n-1} + a_1 r^{n-2} + \ldots + a_{n-2} r + a_{n-1}. \end{aligned}$$

Hiermit ist unsere Behauptung bewiesen, die wir auch so ausdrücken können: Dividiert man $F(x)$ durch $(x - r)$, so ist der Rest gleich $F(r)$.

Wenn $F(x)$ für den Wert $x = r$ verschwindet, so nennt man r eine Wurzel der Gleichung

$$F(x) = 0.$$

Das Problem, zu einer Gleichung alle rationalen Wurzeln zu bestimmen, ist daher gleichbedeutend damit, alle Primfunktionen ersten Grades $(x - r)$ zu ermitteln, durch die $F(x)$ teilbar ist, und umgekehrt. Die im vorigen Paragraphen angegebene allgemeine Methode läfst sich hierbei noch etwas modifizieren und führt zu sehr einfachen Resultaten.

Wenn die Funktion $F(x)$, die wir als ganzzahlig voraussetzen, durch $(x - r)$ teilbar ist, so ist $F(tx)$ durch $(tx - r)$ oder $\left(x - \frac{r}{t}\right)$ teilbar, wenn t eine beliebige Konstante ist. Setzen wir nun, wenn

$$F(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a^{n-1} x + a_n$$

ist, $t = a_0$, $tx = y$, so mufs die ganzzahlige Funktion

$$F_1(y) = y^n + a_1 y^{n-1} + a_0 a_2 y^{n-2} + \ldots + a_0^{n-1} a_n$$

durch eine lineare Primfunktion teilbar sein. Diese mufs nun aber nach den Sätzen am Schlusse von § 57 die Form $(y - r)$ haben, wo r eine ganze Zahl ist. Da $F_1(0) = a_0^{n-1} a_n$ durch die ganze Zahl r teilbar sein mufs, so hat man nur eine beschränkte Anzahl von Werten für r, nämlich jeden Teiler von $a_0^{n-1} a_n$, mit positivem oder negativem Vorzeichen in Betracht zu ziehen und zuzusehen, ob für einen von ihnen $F(r) = 0$ wird.

Einige Beispiele werden die Anwendung dieses Verfahrens klar machen.

1) Die Funktion $x^2 + 1$ ist eine Primfunktion, da sich bei der Division von $x - 1$ und $x + 1$ der Rest 2 ergiebt, oder auch schon deshalb, weil $x^2 + 1$ für keinen reellen Wert verschwindet.

2) Die Funktion $x^2 - a$ ist dann eine Primfunktion, wenn a keine Quadratzahl ist. Ist dagegen $a = r^2$ Quadratzahl, so lautet die Zerlegung $x^2 - a = (x - r)(x + r)$.

3) Um die allgemeine quadratische Funktion

$$ax^2 + 2bx + c$$

zu untersuchen, multiplizieren wir sie mit a und erhalten in

$$a(ax + 2bx + c) = (ax + b)^2 - (b^2 - ac)$$

eine Funktion von der im vorigen Beispiel angegebenen Form. Nur wenn $b^2 - ac > 0$ und eine Quadratzahl ist, ist also $ax^2 + 2bc + c$ reduktibel, in allen übrigen Fällen dagegen eine Primfunktion. Der Zusammenhang mit der in der elementaren Algebra gelehrten Auflösungsmethode der quadratischen Gleichungen (Bd. I, § 27) ist hier sofort ersichtlich.

4) Betrachten wir noch als Beispiel eine Funktion von höherem Grade $F(x) = x^4 + 4x^3 + 6x^2 - x - 10$ und beachten sämtliche Teiler von 10, nämlich $\pm 1, \pm 2, \pm 5, \pm 10$, so erhalten wir sofort $F(1) = 0$, $F(-2) = 0$ und die Zerlegung

$$x^4 + 4x^3 + 6x^2 - x - 10 = (x - 1)(x + 2)(x^2 + 3x + 5),$$

wobei $x^2 + 3x + 5$ leicht als Primfunktion kenntlich ist.

---

§ 60. Irreduktibilität gewisser Funktionen.

Ist die Aufsuchung aller Primfunktionen ersten Grades, die in einer gegebenen Funktion als Teiler enthalten sind, verhältnismäſsig einfach, so wachsen doch die Schwierigkeiten, alle übrigen zu bestimmen, mit zunehmendem Grade beträchtlich. Für die Aufstellung aller in einer gegebenen Funktion enthaltenen Primfunktionen zweiten Grades läſst sich mit Hülfe der Kettenbrüche ein sehr elegantes Verfahren aufstellen, doch würde es uns zu weit führen, diese hier darzulegen, zumal die prinzipielle Seite der Sache durch die Untersuchungen des § 57 hinreichend klargelegt ist.*) Wir wollen uns begnügen, eine bestimmte Art von Funktionen als irreduktibel zu charakterisieren:

*) Es möge verwiesen werden auf Lagrange, Traité de la résolution des équations.

Wenn in einer ganzzahligen Funktion das Produkt des Koeffizienten der höchsten Potenz der Unbestimmten und des von dieser freien Koeffizienten durch eine Primzahl, aber nicht durch deren Quadrat, alle übrigen Koeffizienten aber durch diese Primzahl teilbar sind, so ist die Funktion eine Primfunktion.

Wir bezeichnen diese Funktion mit C(x), zwei komplementäre ganze ganzzahlige Teiler mit A(x) und B(x), so dafs

$$\begin{aligned} A(x) &= a_0 x^\alpha + a_1 x^{\alpha-1} + \ldots + a_\alpha \\ B(x) &= b_0 x^\beta + b_1 x^{\beta-1} + \ldots + b_\beta \\ C(x) &= c_0 x^\gamma + c_1 x^{\gamma-1} + \ldots + c_\gamma \\ A(x)\, B(x) &= C(x), \quad \alpha + \beta = \gamma \end{aligned}$$

sein soll. Da $c_0 c_\gamma$ durch eine Primzahl p, aber nicht durch p^2 teilbar sein soll, so mufs einer der beiden Koeffizienten c_0 oder c_γ dieselbe Eigenschaft haben, während der andere zu p teilerfremd ist. Sei c_0 der letztere, so folgt aus $c_\gamma = a_\alpha b_\beta$, dafs wieder der eine von den beiden Koeffizienten a_α, b_β durch p teilbar ist. Sei a_α dieser, so ist b_β sicher teilerfremd zu p. Alle Koeffizienten von A(x) können nicht durch p teilbar sein, weil sonst $a_0 b_0 = c_0$ es wäre. Wir nehmen also an, dafs a_i der letzte ist, der zu p teilerfremd ist, so dafs also $a_{i+1}, a_{i+2} \ldots a_\alpha$ durch p teilbar seien. Betrachten wir nun den Koeffizienten

$$c_{\beta+i} = a_i b_\beta + a_{i+1} b_{\beta-1} + a_{i+2} b_{\beta-2} + \ldots,$$

so ergiebt sich sofort, dafs dieser dann nicht durch p teilbar ist, da $a_i a_\beta$ es nicht ist, während alle übrigen Glieder teilbar sind. Es mufs daher $\beta = i = 0$ sein, d. h. die Funktion C(x) ist eine Primfunktion.

Diesen Satz wollen wir nun anwenden, um die Funktion $x^p - 1$ zu untersuchen, wo p wieder eine beliebige Primzahl ist. Offenbar hat diese den Teiler $x - 1$; der complementäre Teiler hierzu

$$\varphi_p(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \ldots + x + 1$$

erweist sich nun als Primfunktion; denn bildet man

$$\varphi_p(x+1) = \frac{(x+1)^p - 1}{x} = x^{p-1} + p x^{p-2} + \frac{p(p-1)}{1.2} x^{p-3} + \ldots \frac{p(p-1)}{1.2} x + p,$$

so erhält man eine Funktion von der soeben charakterisierten Art. Daher ist $\varphi_p(x+1)$ und ebenso $\varphi_p(x)$ eine Primfunktion.

---

§ 61. Modulsysteme aus ganzen Funktionen.

Aus dem Begriffe der Teilbarkeit leitet man folgende Fundamentalsätze ab:

Die Summe, die Differenz, allgemein jedes Aggregat von ganzen Funktionen ist durch jeden ihnen gemeinschaftlichen Teiler teilbar.

Ein Produkt von ganzen Funktionen ist durch jeden Teiler seiner Faktoren teilbar.

Diese Fundamentalsätze entsprechen völlig denen, die wir in § 15 für die ganzen Zahlen aufgestellt haben, doch möge hier bemerkt werden, dafs der zweite Satz sich nicht wie dort aus dem ersten ableiten läfst, sondern ihm selbständig und unabhängig gegenübersteht.

Um auszusagen, dafs die gemeinschaftlichen Teiler der ganzen Funktionen

$$A_1(x),\ A_2(x), \ldots A_n(x)$$

vollständig mit denen der Funktionen

$$B_1(x),\ B_2(x), \ldots B_m(x)$$

übereinstimmen, gebrauchen wir von jetzt ab, indem wir die Funktionen zu Systemen zusammenfassen, die kürzere Ausdrucksweise, dafs das aus den ersteren gebildete Modulsystem dem aus den letzteren gebildeten äquivalent ist, was in der Formel

$$[A_1(x), A_2(x), \ldots A_n(x)] = [B_1(x), B_2(x), \ldots B_m(x)]$$

zur Bezeichnung gelangen soll. Um zu entscheiden, ob zwei Modulsysteme äquivalent sind, bestände das begrifflich einfachste Verfahren darin, alle Teiler der einzelnen Funktionen zu bestimmen und dann die gemeinschaftlichen

des einen Systems mit denen des andern zu vergleichen, wobei man sich auf primitive Teiler beschränken könnte, sodafs man durch eine beschränkte Anzahl von Operationen zum Ziele gelangen mufs. Dieses Verfahren ist nun aber höchst umständlich und unzweckmäfsig, weil folgende einfache Gesetze gelten:

Ein Modulsystem bleibt in seiner Äquivalenz ungeändert, wenn man irgend eins seiner Elemente zu einem andern addiert, oder es von diesem subtrahiert.

Ein Modulsystem bleibt sich selbst äquivalent, wenn man ihm ein Produkt aus einer beliebigen ganzen Funktion in eins seiner Elemente hinzufügt.

Den ersten Satz, den wir durch

$$[A(x), B(x), C(x)\ldots)] = [A(x) \pm B(x), B(x), C(x), \ldots]$$

ausdrücken können, beweisen wir in folgender Weise. Jeder gemeinschaftliche Teiler von $A(x)$ und $B(x)$ mufs auch $A(x) \pm B(x)$ teilen, daher ist auch jeder gemeinschaftliche Teiler von $A(x), B(x), C(x)\ldots$ ein gemeinschaftlicher Teiler von $A(x) \pm B(x), B(x), C(x), \ldots$ Umgekehrt mufs auch jeder gemeinschaftlicher Teiler von $A(x) \pm B(x)$ und $B(x)$ ein Teiler von $[A(x) \pm B(x)] \mp B(x) = A(x)$ sein und daher jeder gemeinschaftlicher Teiler von $A(x) \pm B(x), B(x), C(x)\ldots$ auch ein ebensolcher von $A(x), B(x), C(x)\ldots$ sein. Der zweite Satz, dem wir die Form

$$[A(x), B(x), C(x), \ldots] = [G(x)\, A(x), A(x), B(x), C(x), \ldots]$$

geben können, geht einfach daraus hervor, dafs $G(x)\, A(x)$ durch $A(x)$ und jeden Teiler von $A(x)$ teilbar ist. Wendet man nun diese Sätze wiederholt an, so erhält man leicht die Formeln

$$[A_0(x), A_1(x), A_2(x), \ldots A_n(x)] = [A_0(x) + G_1(x)\, A_1(x) + G_2(x)\, A_2(x) + \ldots + G_n(x)\, A_n(x), A_1(x), A_2(x), \ldots A_n(x)]$$

$$[A_1(x), A_2(x), \ldots A_n(x)] = [G_1(x)\, A_1(x) + G_2(x)\, A_2(x) + \ldots + G_n(x)\, A_n(x), A_1(x), A_2(x), \ldots A_n(x)],$$

deren Inhalt durch die folgenden Sätze wiedergegeben wird:

Ein Modulsystem aus ganzen Funktionen wird in seiner Äquivalenz nicht geändert, wenn man zu

irgend einem seiner Elemente eine lineare Form der übrigen mit beliebigen ganzen Funktionen als Koeffizienten addiert.

Einem aus ganzen Funktionen als Elementen bestehenden Modulsystem kann man unbeschadet der Äquivalenz eine lineare Form der Elemente mit ganzen Funktionen als Koeffizienten als neues Element beifügen.

Ganz allgemein findet man:

Wenn die Elemente zweier Modulsysteme in einem solchen Zusammenhang zu einander stehen, dafs alle Elemente des einen sich als lineare Formen des andern ausdrücken lassen und umgekehrt, so sind die Modulsysteme äquivalent.

Denn haben die beiden Modulsysteme

$$[A_1(x), A_2(x), \ldots A_n(x)]$$

und

$$[B_1(x), B_2(x), \ldots B_m(x)]$$

diese Eigenschaft, so ist nach den voraufgehenden Sätzen

$$[A_1(x), A_2(x), \ldots A_n(x)]$$
$$= [A_1(x), A_2(x), \ldots A_n(x), B_1(x), B_2(x), \ldots B_m(x)]$$
$$[B_1(x), B_2(x), \ldots B_m(x)]$$
$$= [B_1(x), B_2(x), \ldots B_m(x), A_1(x), A_2(x), \ldots A_n(x)]$$

und folglich

$$[A_1(x), A_2(x), \ldots A_n(x)] = [B_1(x), B_2(x), \ldots B_m(x)],$$

ebenso wie wir das in § 17 bei Modulsystemen aus ganzen Zahlen gesehen haben.

---

§ 62. Reduktion der Modulsysteme. Gröfster gemeinschaftlicher Teiler.

Die bisher über Modulsysteme abgeleiteten Sätze können nun dazu benutzt werden, um ein Modulsystem zu reduzieren und zwar schliefslich soweit, dafs es nur eine einzige ganze Funktion enthält.

Betrachten wir zunächst den Fall, dafs das Modulsystem nur zwei ganze Funktionen A(x) und B(x) enthält,

deren Grade α und β seien, so dafs etwa $\alpha > \beta$ angenommen wird. Wie wir in § 56 gezeigt haben, kann man immer zwei ganze Funktionen G(x) und C(x), die letztere von einem Grade $\gamma < \beta$, so bestimmen, dafs

$$A(x) = G(x) B(x) + C(x)$$

ist. Dann ergiebt sich aber sofort, dafs

$$[A(x), B(x)] = [B(x), C(x)]$$

ist, womit eine Reduktion erzielt ist, da das rechts stehende System Funktionen von niederem Grade enthält als das links stehende. Ist nun $C(x) \neq 0$, so kann man auf die beiden Funktionen B(x) und C(x) dasselbe Verfahren anwenden, wie auf A(x) und B(x), und allgemein zu A(x), B(x) eine Reihe von neuen Funktionen C(x), D(x), ... bilden, von denen immer zwei aufeinanderfolgende äquivalente Modulsysteme ergeben, und deren Grade beständig abnehmen. Aus dem letzteren Grunde mufs schliefslich die Reihe abbrechen und mit einer Null schliefsen. Nennt man das voraufgehende Glied der Reihe T(x), so ist das letzte Modulsystem $[T(x), 0] = T(x)$ das reduzierte und stellt daher den gröfsten gemeinschaftlichen Teiler von A(x) und B(x) dar, da jeder andere in ihm enthalten sein mufs. Ist T(x) eine Konstante, so haben A(x) und B(x) keine eigentliche ganze Funktion mehr als Teiler und werden dann teilerfremd oder relativ prim genannt.

Ganz ähnlich wie bei einem aus zwei Elementen bestehenden Modulsystem ist nun auch das Verfahren bei mehreren Elementen, wobei man immer schrittweise eine Funktion durch eine von niederem Grade ersetzen kann, bis schliefslich nur noch eine einzige nicht verschwindende Funktion vorhanden ist, die den gröfsten gemeinschaftlichen Teiler aller Elemente des Modulsystems darstellt, und deren Koeffizienten dem Rationalitätsbereich der Koeffizienten der einzelnen Funktionen angehören.

Achtet man nun auf die Abhängigkeit der Elemente zweier bei einer solchen Reduktion auf einanderfolgender Modalsysteme, so findet man, dafs sie durch eine lineare Transformation mit ganzen Funktionen als Koeffizienten auseinander hervorgehen, und gelangt dann durch dieselben Schlüsse wie in § 18 zu dem Satze:

Die Elemente zweier äquivalenter Modulsysteme sind wechselseitig durch einander als lineare Formen mit ganzen Funktionen als Koeffizienten darstellbar.

Wenn man dann noch den Schlufssatz im vorigen Paragraphen beachtet, so ergiebt sich:

Die hinreichende und notwendige Bedingung dafür, dafs zwei aus ganzen Funktionen bestehenden Modulsysteme äquivalent sind, besteht darin, dafs sich ihre Elemente wechselseitig linear und homogen durcheinander darstellen lassen.

Jetzt lassen sich alle Betrachtungen, die wir früher über Modulsysteme mit ganzen Zahlen in § 19 und 20 angestellt haben, ohne Weiteres auf Modulsysteme, die aus ganzen Funktionen bestehen, zur Anwendung bringen. Hervorheben wollen wir nur den Satz, dafs jede ganze Funktion sich nur auf eine einzige Weise als ein Produkt von lauter Primfunktionen darstellen läfst. Wenn man diese Zerlegung kennt, so kann man die sämtlichen eigentlichen Teiler einer ganzen Funktion, den gröfsten gemeinschaftlichen Teiler und das kleinste gemeinschaftliche Vielfache von mehreren ganzen Funktionen genau so bestimmen, wie wir es bei den ganzen Zahlen dargelegt haben. (§ 21).

Da die ganzen linearen Funktionen stets Primfunktionen sind, so mufs eine ganze Funktion, wenn sie durch lauter verschiedene solche Funktionen $x - a_1$, $x - a_2, \ldots x - a_n$ teilbar ist, auch durch deren Produkt teilbar sein. Die Anzahl solcher Linearfaktoren kann also niemals den Grad der Funktion überschreiten. Eine Gleichung

$$F(x) = 0$$

vom nten Grade kann also niemals mehr als n Wurzeln besitzen, da jedem Linearfaktor von $F(x)$ eine Wurzel der Gleichung entspricht. Hieraus folgt unmittelbar: Wenn zwei ganze Funktionen vom nten Grade für mehr als n Werte der Variabeln übereinstimmen, so stimmen sie in den Koeffizienten gleicher Potenzen der Variabeln überein.

Denn bildet man die Differenz der beiden Funktionen, so würde diese sonst eine ganze Funktion von n tem oder niederem Grade darstellen, die mehr als n Wurzeln hätte. Man sagt von zwei ganzen Funktionen, die in den Koeffizienten gleicher Potenzen der Variabeln übereinstimmen, auch, daſs sie identisch gleich sind.

Wenn eine ganze Funktion nicht identisch verschwindet, so verschwindet sie auch für unendlich viele Werte nicht, die man an Stelle der Veränderlichen setzen kann. Das gilt auch noch für ein System von ganzen Funktionen und läſst sich ferner auf Funktionen von mehreren Variablen verallgemeinern: Verschwinden die Funktionen eines solchen Systems nicht für alle Werte der Variabeln, so kann man diesen unendlich viele Wertsysteme beilegen, so daſs für sie keine der Funktionen den Wert Null annimmt. Ist der Beweis für den Fall erbracht, daſs die Funktionen von n Unbestimmten abhängen, so läſst sich zeigen, daſs die Behauptung auch für $(n+1)$ Variable und somit allgemein gültig ist. Denken wir uns nämlich alle Funktionen eines solchen Systems nach Potenzen einer Variabeln entwickelt, so bilden die Koeffizienten ein Funktionensystem, das von den n übrigen Variabeln abhängt. Man kann diesen also solche numerischen Werte geben, daſs alle Koeffizienten nicht verschwinden oder wenigstens nicht alle solche, die einer und derselben Funktion angehören. Wenn das geschehen ist, so sind für die letzte Variable noch unendlich viele Werte möglich, die der gestellten Forderung Genüge leisten.

---

§ 63. Beispiele.

Wir wollen jetzt an einigen Beispielen die Methode der Aufsuchung des gröſsten gemeinschaftlichen Teilers erläutern.

1) Für zwei lineare Funktionen

$$A(x) = a_0 x + a_1, \quad B(x) = b_0 x + b_1, \quad (a_0 \neq 0, b_0 \neq 0)$$

erhält man aus

$$A(x) = \frac{a_0}{b_0} B(x) + \frac{a_1 b_0 - a_0 b_1}{b_0},$$

daſs

$$[A(x), B(x)] = [B(x), a_0 b_1 - a_1 b_0]$$

ist. A(x) und B(x) sind teilerfremd oder gegenseitig durch einander teilbar, je nachdem die Determinante $a_0 b_1 - a_1 b_0$ von Null verschieden ist oder verschwindet.

2) Sei

$$A(x) = a_0 x^2 + a_1 x + a_2,\quad B(x) = b_0 x + b_1,\quad (a_0 \neq 0;\ b_0 \neq 0)$$

so wird

$$A(x) = G(x)\, B(x) + C,$$

wenn

$$G(x) = g_0 x + g_1,\quad g_0 = \frac{a_0}{b_0},\quad g_1 = \frac{a_1 b_0 - a_0 b_1}{b_0^2},$$

$$C = \frac{a_0 b_1^2 - a_1 b_1 b_0 + a_2 b_0^2}{b_0^2}$$

gesetzt wird. Es ist daher B(x) ein Teiler von A(x) oder nicht, je nachdem

$$a_0 b_1^2 - a_1 b_1 b_0 + a_2 b_0^2$$

verschwindet oder nicht.

3) Nehmen wir

$$A(x) = a_0 x^2 + a_1 x + a_2,\quad B(x) = b_0 x^2 + b_1 x + b_2$$
$$(a_0 \neq 0;\ b_0 \neq 0)$$

an, so erhalten wir

$$A(x) = \frac{a_0}{b_0} B(x) + C(x),$$

wo

$$C(x) = c_0 x + c_1,\quad c_0 = \frac{a_1 b_0 - a_0 b_1}{b_0},\quad c_1 = \frac{a_2 b_0 - a_0 b_2}{b_0}.$$

Ist nun $c_0 = 0$ oder statt dessen die Determinante

$$a_0 b_1 - a_1 b_0 = 0,$$

so sind A(x) und B(x) gegenseitig durch einander teilbar oder teilerfremd, je nachdem die Determinante $a_2 b_0 - a_0 b_2$ verschwindet oder nicht. Ist aber $c_0 \neq 0$, so kann man weiter bilden

$$B(x) = G(x)\, C(x) + D,$$

wo

$$G(x) = g_0 x + g_1,\quad g_0 = \frac{b_0}{c_0},\quad g_1 = \frac{b_1 c_0 - b_0 c_1}{c_0^2},$$

$$D = \frac{b_0 c_1^2 - b_1 c_1 c_0 + b_2 c_0^2}{c_0^2}$$

$$= \frac{(a_0 b_2 - a_2 b_0)^2 - (a_1 b_0 - a_0 b_1)(a_2 b_1 - a_1 b_2)}{c_0^2 b_0}$$

ist. $A(x)$ und $B(x)$ sind also durch $C(x)$ teilbar oder teilerfremd, je nachdem

$$(a_0 b_2 - a_2 b_0)^2 - (a_1 b_0 - a_0 b_1)(a_2 b_1 - a_1 b_2)$$

verschwindet oder nicht.

Fassen wir alles zusammen, so finden wir, dafs $A(x)$ und $B(x)$ dann und nur dann teilerfremd sind, wenn dieser letztere Ausdruck nicht verschwindet, dafs ferner sie einen Teiler vom ersten Grade als gröfsten gemeinschaftlichen Teiler haben oder gegenseitig durcheinander teilbar sind, je nachdem die Determinanten zweiter Ordnung der Matrix

$$\begin{pmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{pmatrix}$$

nicht alle verschwinden oder verschwinden, d. h. diese Matrix den Rang 2 oder 1 hat.

4) Betrachten wir ferner die beiden Funktionen

$$\begin{aligned} A(x) &= a_0 x^3 + a_1 x^2 + a_2 x + a_3 \\ A'(x) &= 3 a_0 x^2 + 2 a_1 x + a_2 \end{aligned} \qquad (a_0 \neq 0)$$

und bilden

$$A(x) = G(x) A'(x) + B(x),$$

wo

$$B(x) = b_0 x + b_1, \quad b_0 = \frac{6 a_0 a_2 - 2 a_1^2}{9 a_0}, \quad b_1 = \frac{9 a_0 a_3 - a_1 a_2}{9 a_0},$$

$$G(x) = g_0 x + g_1, \quad g_0 = \frac{1}{3}, \qquad g_1 = \frac{a_1}{9 a_0}$$

gesetzt ist. Ist nun $b_0 = 0$, so ist $A(x)$ durch $A'(x)$ teilbar, oder es sind beide teilerfremd, je nachdem b_1 verschwindet oder nicht. Im ersteren Falle ist also

$$3 a_0 a_2 - 2 a_1^2 = 0, \quad 9 a_0 a_3 - a_1 a_2 = 0.$$

Ist $b_0 \neq 0$, so bilden wir weiter

$$A'(x) = H(x) B(x) + C,$$

wo

$$H(x) = h_0 x + h_1, \quad h_0 = \frac{3 a_0}{b_0}, \quad h_1 = \frac{2 a_1 b_0 - 3 a_0 b_1}{b_0^2},$$

$$C = \frac{3 a_0 b_1^2 - 2 a_1 b_1 b_0 + a_2 b_0^2}{b_0^2}$$

$$= \frac{27 a_0 a_3^2 + 4 a_0 a_2^3 + 4 a_1^3 a_2 - a_1^2 a_2^2 - 18 a_0 a_1 a_2 a_3}{9 b_0^2}$$

ist. In diesem Falle haben $A(x)$ und $A'(x)$ einen Teiler ersten Grades oder sind teilerfremd, je nachdem

$$27 a_0 a_3^2 + 4 a_0 a_2^3 + 4 a_1^3 a_2 - a_1^2 a_2^2 - 18 a_0 a_1 a_2 a_3$$

verschwindet oder nicht.

Faſst man alles zusammen, so zeigt sich, daſs $A(x)$ und $A'(x)$ dann und nur dann teilerfremd sind, wenn dieser letztere Ausdruck nicht verschwindet, daſs bei seinem Verschwinden dagegen der Rang der Matrix

$$\begin{pmatrix} 3 a_0 & a_1 & a_2 \\ a_1 & a_2 & 3 a_3 \end{pmatrix}$$

entscheidet, ob $A(x)$ und $A'(x)$ einen bloſsen Linearfaktor gemeinsam haben, oder ob $A(x)$ durch $A'(x)$ teilbar ist; das Erstere findet statt, wenn der Rang gleich 2, das Letztere, wenn er gleich 1 ist.

5) Sind a und b ganze positive Zahlen, so ist

$$(x^a - 1, \; x^b - 1) = x^{(a, b)} - 1.$$

Dies läſst sich in folgender Weise zeigen. Ist $a > b$ und bestimmt man g und c so, daſs

$$a = g b + c, \; c < b,$$

so ist

$$x^a - 1 = (x^{g b} - 1) x^c + x^c - 1,$$

und es läſst sich $x^{g b} - 1$ darstellen in der Form

$$x^{b g} - 1 = (x^b)^g - 1 = (x^b - 1)(x^{b g - b} + \ldots + x^b + 1),$$

so daſs

$$(x^a - 1, x^b - 1) = (x^b - 1, x^c - 1)$$

wird. So kann man fortfahren, und es ist leicht ersichtlich, daſs die in der Reihe der Funktionen

$$x^a - 1, \; x^b - 1, \; x^c - 1, \ldots$$

auftretenden Exponenten a, b, c, ... dieselben Zahlen sind, die bei der Bestimmung des gröſsten gemeinschaftlichen Teilers von a und b sich zeigen, von denen die beiden letzten $t = (a, b)$ und 0 sind. Hieraus ergiebt sich sofort das angegebene Resultat.

Bezeichnet man den Quotienten $\frac{x^m - 1}{x - 1}$ mit $\varphi_m(x)$, so ist ebenso

$$[\varphi_a(x), \varphi_b(x)] = \varphi_{(a, b)}(x).$$

Dies läſst sich natürlich auch direkt beweisen, indem man beachtet, daſs für $a > b$

$$\varphi_a(x) = x^{a-b}\varphi_b(x) + \varphi_{a-b}(x)$$

ist u. s. f.

6) Aus dem Vorherigen ergiebt sich leicht:

$$(x^{a_1} - 1, \; x^{a_2} - 1, \ldots x^{a_n} - 1) = x^{(a_1, a_2, \ldots a_n)} - 1.$$

Ferner läſst sich ableiten, daſs

$$(x^{p^{a_1}} - x, \; x^{p^{a_2}} - x, \ldots x^{p^{a_n}} - x) = x^{p^{(a_1, a_2, \ldots a_n)}} - x$$

ist.

Weitere Beispiele mit numerischen Koeffizienten kann sich der Leser selbst leicht bilden.

---

§ 64. Zerlegung ganzer ganzzahliger Funktionen von mehreren Variabeln.

Eine ganze Funktion einer Unbestimmten x mit Koeffizienten eines beliebigen Rationalitätsbereiches läſst sich immer nur auf eine einzige Weise in lauter Primfunktionen zerlegen, deren Koeffizienten demselben Rationalitätsbereich angehören. Ist nun der Rationalitätsbereich ein natürlicher, so ergiebt sich aus den Sätzen in § 57, daſs sich jede ganze ganzzahlige Funktion nur auf eine einzige Weise als ein Produkt des wesentlichen Zahlenteilers ihrer Koeffizienten und lauter solcher Primfunktionen zerlegen läſst, die selbst auch keinen Zahlenteiler mehr haben und also nach unserer früheren Bezeichnungsweise primitiv sind.

Dieser Satz läſst sich nun auf ganze ganzzahlige Funktionen von beliebig vielen Veränderlichen durch schrittweise Verallgemeinerung ausdehnen. Wir wollen annehmen, daſs er für ganze ganzzahlige Funktionen von n Veränderlichen $x_1, x_2, \ldots x_n$ schon bewiesen sei und darauf zeigen, daſs er auch für ebensolche Funktionen von $(n+1)$ Veränderlichen $x, x_1, x_2, \ldots x_n$ gilt. Betrachten wir jedoch zunächst die ganzen ganzzahligen Funktionen von n Veränderlichen etwas genauer, so finden wir daſs bei ihnen auſser Primzahlen Primfunktionen von n verschiedenen Arten als Teiler auftreten können, nämlich ganze ganzzahlige Primfunktionen einer der Variabeln, ebensolche von zwei der Variabeln u. s. w., solche von $(n-1)$ Variabeln und schlieſslich Primfunktionen von allen n Variabeln. Die ganzen ganzzahligen Primfunktionen von k der Variabeln können wir als Primfunktionen kter Stufe bezeichnen, wo k die Werte $1, 2, \ldots n$ annehmen darf, und wenn wir wollen, so können wir diese Bezeichnungsweise auch auf die Primzahlen ausdehnen, die dann als Primfunktionen nullter Stufe angesehen werden können. Alle diese Behauptungen wollen wir also als bewiesen annehmen für den Fall, daſs die Anzahl der Variabeln gleich n ist, und darauf zeigen, daſs sie auch noch gelten, wenn eine neue Variable x zu den bisherigen $x_1, x_2, \ldots x_n$ hinzutritt. Dies geschieht, indem wir die Sätze in § 57 verallgemeinern.

Es seien $A(x)$ und $B(x)$ ganze ganzzahlige Funktionen von $x, x_1, x_2, \ldots x_n$, ihr Produkt $C(x)$ ist dann ebenfalls eine solche Funktion. Stellen wir die Funktionen wie in § 57 dar in der Form

$$A(x) = a_0 x^\alpha + a_1 x^{\alpha-1} + \ldots + a_{\alpha-1} x + a_\alpha$$
$$B(x) = b_0 x^\beta + b_1 x^{\beta-1} + \ldots + b_{\beta-1} x + b_\beta$$
$$C(x) = c_0 x^\gamma + c_1 x^{\gamma-1} + \ldots + c_{\gamma-1} x + c_\gamma,$$

wo $\gamma = \alpha + \beta$ ist, so sind die Koeffizienten ganze ganzzahlige Funktionen von $x_1, x_2, \ldots x_n$, zwischen denen die Gleichungen

$$c_h = \Sigma\, a_i b_k$$

bestehen, wo die Summe rechts über alle Koeffizienten a_i, b_k zu erstrecken ist, deren Indizes der Bedingung $i + k = h$ genügen.

Die gröfsten gemeinschaftlichen Teiler der Koeffizienten

$$\mathfrak{a} = (a_0, a_1, \ldots a_\alpha)$$
$$\mathfrak{b} = (b_0, b_1, \ldots b_\beta)$$
$$\mathfrak{c} = (c_0, c_1, \ldots c_\gamma)$$

sind ebenfalls ganze ganzzahlige Funktionen von $x_1, x_2 \ldots x_n$. Nun lassen sich die Betrachtungen des § 57 Wort für Wort auf die jetzt vorliegenden Verhältnisse übertragen, nur dafs man für p, das dort als eine Primzahl angenommen wurde, jetzt nicht nur solche, sondern auch beliebige Primfunktionen erster bis n ter Stufe setzen kann. Es ergiebt sich dann ebenso wie dort, dafs

$$\mathfrak{a}\,\mathfrak{b} = \mathfrak{c}$$

ist, dafs also das Produkt der Teiler gleich dem Teiler des Produktes ist, falls man als Teiler die natürlichen Zahlen, sowie die Funktionen von erster bis n ter Stufe in Betracht zieht. Hieraus läfst sich nun weiter wie früher ableiten, dafs wenn eine ganze ganzahlige Funktion von $x, x_1, \ldots x_n$ sich in ein Produkt von lauter ganzen Funktionen von x mit Koeffizienten des Rationalitätsbereiches $(x_1, x_2, \ldots x_n)$ zerlegen läfst, man die Faktoren des Produktes auch als ganze ganzzahlige Funktionen von $x, x_1, x_2, \ldots x_n$ herstellen kann.

Nun läfst sich aber jede ganze Funktion von x mit Koeffizienten des Rationalitätsbereiches $(x_1, x_2, \ldots x_n)$ nur auf eine Weise als ein Produkt von Primfunktionen von x zerlegen, wie aus § 62 hervorgeht. Hierbei sind aber zunächst die Gröfsen des Rationalitätsbereiches nicht mit in Betracht gezogen. Nimmt man aber auch auf sie Rücksicht, so kann man jeder auftretenden Primfunktion eine solche eindeutig entsprechen lassen, deren Koeffizienten ganz und ganzzahlig in $x_1, x_2, \ldots x_n$ sind und keinen Zahlenteiler oder Teiler von erster bis n ter Stufe gemeinsam haben. In Bezug auf diese Primfunktionen ist die Zerlegung ebenfalls eindeutig. Wenn nun aber eine ganze ganzzahlige Funktion von $x, x_1, x_2, \ldots x_n$ vorliegt, so mufs, wenn man alle solche Primfunktionen ausscheidet, der überschüssige Faktor ganz und ganzzahlig in $x_1, x_2, \ldots x_n$ sein und kann nach unserer Voraussetzung nur auf eine einzige Weise in ganze ganzzahlige Primfunktion zerlegt werden, und so

ergiebt sich dann, dafs auch die ganzen ganzzahligen Funktionen von $(n+1)$ Variabeln $x, x_1, x_2, \ldots x_n$ nur auf eine einzige Weise zerlegbar sind in ein Produkt von Primfunktionen der verschiedenenen Stufen. Die Primfunktionen $P(x, x_1, x_2, \ldots x_n)$ $(n+1)$ter Stufe können übrigens in ähnlicher Weise ermittelt werden, wie wir das in § 58 für die ganzen Funktionen von x im natürlichen Rationalitätsbereich dargelegt haben. Wir wollen dies hier aber nicht ausführlich entwickeln, sondern dem Leser überlassen, die notwendigen Modifikationen in den dort angestellten Betrachtungen selbst vorzunehmen.

---

§ 65. Absonderung mehrfacher Faktoren aus ganzen Funktionen.

Angesichts der grofsen Schwierigkeiten, die das Problem der Zerlegung einer ganzen Funktion in ein Produkt von Primfunktionen darbietet, ist es von grofser Wichtigkeit, dafs man es bedeutend vereinfachen kann, wenn die Funktion Primfunktionen in höheren Potenzen enthält. Es läfst sich dann das Problem auf die Zerlegung solcher ganzer Funktionen zurückführen, die diese Primfunktionen nur einfach enthalten. Man gelangt zu diesem Verfahren, wenn man neben den Funktionen ihre Ableitungen in die Betrachtung einführt. Es ergeben sich hierbei folgende Sätze:

1. Eine ganze Funktion, die zu ihrer Ableitung teilerfremd ist, kann nur in Faktoren zerlegt werden, die zu einander teilerfremd sind.

Nimmt man eine Zerlegung von der Form

$$F(x) = A(x)\,B(x)$$

an, so ist (§ 6)

$$F'(x) = A'(x)\,B(x) + A(x)\,B'(x).$$

Würden nun $A(x)$ und $B(x)$ einen gemeinschaftlichen Teiler haben, so würde dieser auch ein Teiler von $F'(x)$ sein, und da er in $F(x)$ quadratisch vorkommt, so würden $F(x)$ und $F'(x)$ nicht teilerfremd sein.

2. Zerlegt man eine ganze Funktion

$$F(x) = A(x)\,B(x)$$

in ein Produkt von zwei teilerfremden Faktoren $A(x)$ und $B(x)$, so ist

$$[F(x),\ F'(x)] = [A(x),\ A'(x)]\,[B(x),\ B'(x)].$$

Denn fügt man dem Modulsystem

$$[F(x),\ F'(x)] = [A(x)\,B(x),\ A'(x)\,B(x) + B'(x)\,A(x)]$$

einmal $A(x)$, das andere Mal $B(x)$ hinzu, so erhält man nach gehöriger Reduktion auf beiden Seiten, wenn man $[A(x),\ B(x)] = 1$ beachtet,

$$[A(x),\ F'(x)] = [A(x),\ A'(x)\,B(x)]$$
$$= [A(x),\ A'(x)\,A(x),\ A'(x)\,B(x)] = [A(x),\ A'(x)]$$
$$[B(x),\ F'(x)] = [B(x),\ B'(x)\,A(x)] = [B(x),\ B'(x)]$$

und dann durch Multiplikation

$$[A(x),\ F'(x)]\,[B((x),\ F'(x)] = [A(x),\ A'(x)]\,[B(x),\ B'(x)].$$

Das Modulsystem der linken Seite wird aber

$$[A(x)\,B(x),\ F'(x)\,[A(x),\ B(x)],\ F'^2(x)] = [F(x),\ F'(x)],$$

womit der Beweis geliefert ist.

Man kann diesen Satz auf mehrere Faktoren ausdehnen, sodafs wenn eine Zerlegung

$$F(x) = A_1(x)\ A_2(x)\ldots A_n(x)$$

gegeben ist, in der immer je zwei der Faktoren $A(x)$ teilerfremd zu einander sind,

$$[F(x),\ F'(x)] = \prod_i [A_i(x),\ A_i'(x)] \qquad (i = 1, 2 \ldots n)$$

ist.

Nun kann man aber als solche Faktoren Potenzen von Primfunktionen zu Grunde legen. Es erübrigt also noch zu untersuchen, wie diese sich zu ihren Ableitungen verhalten. Das ist nun aber äufserst einfach. Ist nämlich $P(x)$ eine Primfunktion, so folgt aus

$$F(x) = P^e(x)$$

die Ableitung

$$F'(x) = e\,P^{e-1}(x)\ P'(x).$$

Da Zahlenfaktoren nicht in Betracht gezogen werden und das System

$$[P(x), P'(x)] = 1$$

ist, weil $P'(x)$ von niederem Grade ist als $P(x)$ und daher nicht durch $P(x)$ teilbar sein kann, so ergiebt sich:

3. Ist $F(x)$ eine Potenz einer Primfunktion $P(x)$

$$F(x) = P^e(x),$$

so hat man

$$[F(x), F'(x)] = P(x)^{e-1}.$$

4. Wenden wir die Sätze 2) und 3) an auf die Funktion

$$F(x) = P_1^{e_1}(x) P_2^{e_2}(x) \dots P_n^{e_n}(x)$$

an, so erhalten wir, wenn

$$[F(x), F'(x)] = F_1(x)$$
$$F(x) = F_1(x) G(x)$$

gesetzt wird,

$$G(x) = P_1(x) P_2(x) \dots P_n(x),$$

so dafs nach 2)

$$[G(x), G'(x)] = 1$$

ist. Wir können also jede Funktion $F(x)$ in ein Produkt

$$F(x) = G(x) F_1(x)$$

zerlegen, in dem der erste Faktor $G(x)$ nur in einfache Primfunktionen zerlegt werden kann oder zu seiner Ableitung teilerfremd ist.

5. Dieses Verfahren kann nun fortgesetzt werden. Ist nämlich $F_1(x)$ nicht konstant, so kann man von $F_1(x)$ und $F_1'(x)$ den gröfsten gemeinschaftlichen Teiler $F_2(x)$ bilden u. s. w. Fährt man aber in dieser Weise fort, so mufs man, da die Grade der Funktionen $F(x)$, $F_1(x)$, $F_2(x)$, ... immer niedriger werden, zuletzt auf eine Funktion $F_n(x)$ stofsen, die zu ihrer Ableitung teilerfremd ist. Wir haben dann

$$\begin{aligned}
&[F(x), F'(x)] = F_1(x)\\
&[F_1(x), F_1'(x)] = F_2(x)\\
&\quad\vdots\\
&[F_{n-1}(x), F'_{n-1}(x)] = F_n(x)\\
&[F_n(x), F_n'(x)] = 1.
\end{aligned}$$

Setzt man nun

$$F(x) = G(x)\, F_1(x)$$
$$F_1(x) = G_1(x)\, F_2(x)$$
$$\vdots$$
$$F_{n-1}(x) = G_{n-1}(x)\, F_n(x)$$

und der Gleichförmigkeit wegen noch

$$F_n(x) = G_n(x),$$

so resultiert folgende Zerlegung von $F(x)$

$$F(x) = G(x)\, G_1(x) \ldots G_{n-1}(x)\, G_n(x)$$

in ein Produkt von lauter Funktionen, deren Ableitungen mit ihnen selbst keinen gemeinschaftlichen Teiler haben, und von denen jede folgende in der vorhergehenden als Teiler enthalten ist. Nämlich da

$$F_i'(x) = F'_{i+1}(x)\, G_i(x) + F_{i+1}(x)\, G_i'(x) \qquad (i = 0, 1, 2, \ldots n-1)$$

durch $F_{i+1}(x)$ teilbar ist, so mufs es auch $F'_{i+1}\, G_i(x)$ sein. Da aber F_{i+1} und F'_{i+1} den gröfsten gemeinschaftlichen Teiler F_{i+2} haben, so ist G_i durch $\frac{F_{i+1}}{F_{i+2}} = G_{i+1}$ teilbar.

Man kann daher auch

$$G_i(x) = A_i(x)\, G_{i+1}(x) \qquad (i = 0, 1 \ldots n-1)$$
$$G_n(x) = A_n(x)$$

setzen, wo dann die Funktionen $A_i(x)$ zu $A_i'(x)$ ebenfalls wie zu $G_{i+1}(x)$ teilerfremd sind. Es wird dann

$$G_i(x) = A_i(x)\, A_{i+1}(x) \ldots A_n(x) \qquad (i = 0, 1 \ldots n)$$

$$A_{i-1}(x) = \frac{F_{i-1}(x)\, F_{i+1}(x)}{F_i^2(x)}$$

$$F_i(x) = A_i(x)\, A_{i+1}^2(x) \ldots A_n^{n-i}(x) \quad (i=0, 1 \ldots n-1)$$

und insbesondere

$$F(x) = A_0(x)\, A_1^2(x)\, A_2^3(x) \ldots A_n^{n+1}(x).$$

Das Problem der Zerlegung von $F(x)$ in ein Produkt von Primfunktionen ist damit zurückgeführt auf das einfachere, jede der Funktionen $A_i(x)$ zu zerlegen und zwar in ein Produkt von nur einfach auftretenden Primfunktionen.

---

§ 66. Zerlegung gebrochener Funktionen in Partialbrüche.

Die Betrachtungen, die wir im vierten Abschnitt über lineare Modulsysteme mit ganzzahligen Koeffizienten angestellt haben, lassen sich sehr leicht auf den Fall verallgemeinern, dafs die Koeffizienten ganze Funktionen einer unbestimmten Veränderlichen sind. Man gelangt hierbei zu folgenden Sätzen, bei deren Formulierung wir durch grofse Buchstaben Funktionen von x bezeichnen, dieses Argument aber weglassen.

1) Jedes lineare Modulsystem von der Form

$$(A_1 X + B_1, A_2 X + B_2, \ldots A_n X + B_n)$$

läfst sich auf die Form

$$(AX + B, C)$$

reduzieren, wo

$$A = (A_1, A_2, \ldots A_n)$$
$$(B, C) = (B_1, B_2, \ldots B_n)$$
$$AC = (\ldots A_i B_k - A_k B_k \ldots) \qquad (i, k = 1, 2, \ldots n)$$

ist.

2) Die hinreichende und notwendige Bedingung für die Lösbarkeit der Kongruenz

$$AX + B \equiv 0 \text{ mod } M$$

lautet

$$(A, B, M) = (A, B),$$

und zwar sind die Lösungen eindeutig in Bezug auf $\frac{M}{(A, B)}$ als Modul bestimmt.

3) Sind $M_1, M_2, \ldots M_n$ n Funktionen von x, von denen immer je zwei zu einander teilerfremd sind, so kann man immer in Bezug auf ihr Produkt

$$M = M_1 M_2 \ldots M_n$$

als Modul eindeutig eine Funktion bestimmen, die in Bezug auf $M_1, M_2, \ldots M_n$ gegebene Reste $A_1, A_2 \ldots A_n$ hat, die also den Kongruenzen

$$X \equiv A_i \text{ mod } M_i \qquad (i = 1, 2, \ldots n)$$

genügt. Diese Funktion ist durch eine Kongruenz von der Form

$$X \equiv \sum_i A_i H_i \mod M, \qquad (i = 1, 2, \dots n)$$

bestimmt. Setzt man

$$\begin{aligned} M &= L_i M_i \\ H_i &= L_i X_i, \end{aligned} \qquad (i = 1, 2, \dots n)$$

so genügt X_i der Kongruenz

$$L_i X_i \equiv 1 \mod M_i$$

und ist durch sie eindeutig nach dem Modul M_i bestimmt.

4) Betrachten wir nun irgend eine gebrochene rationale Funktion $\frac{F(x)}{M(x)}$, zerlegen den Nenner in lauter Faktoren $M_1, M_2 \dots M_n$, die zu einander teilerfremd sind; und bestimmen die Reste $A_1, A_2 \dots A_n$ von $F(x)$ nach diesen Faktoren als Moduln, so können wir

$$F(x) \equiv \sum_i A_i L_i X_i \mod M \qquad (i = 1, 2 \dots n)$$

setzen, wo die Funktion X_i nach M_i als Modul völlig bestimmt sind. Diese Kongruenz können wir als Gleichung schreiben

$$F(x) = G_0(x) M(x) + \sum_i A_i L_i X_i, \qquad (i = 1, 2 \dots n)$$

wo $G_0(x)$ eine ganze Funktion bedeutet, und wenn wir dann beide Seiten durch $M(x)$ dividieren, so erhalten wir

$$\frac{F(x)}{M(x)} = G_0(x) + \sum_i \frac{A_i(x) X_i(x)}{M_i(x)}. \qquad (i = 1, 2 \dots n)$$

Ist nun der Grad von $A_i X_i$ größer als der von M_i, so kann man

$$\frac{A_i X_i}{M_i} = G_i + \frac{F_i}{M_i} \qquad (i = 1, 2 \dots n)$$

setzen, wo G_i und F_i ganze Funktionen sind, von denen die letztere einen geringeren Grad hat als $M_i(x)$. Führen wir endlich noch

$$G(x) = G_0(x) + G_1(x) + \dots + G_n(x)$$

ein, so erhalten wir eine Darstellung von folgender Form

$$\frac{F(x)}{M(x)} = G(x) + \sum_i \frac{F_i(x)}{M_i(x)}, \qquad (i = 1, 2, \dots n)$$

bei der die Funktionen $G(x)$ und $F_i(x)$ völlig bestimmt sind. Es läfst sich also jede gebrochene rationale Funktion in bestimmter Weise als eine Summe von einer ganzen Funktion und lauter echt gebrochenen Funktionen darstellen, deren Nenner zu einander teilerfremde Faktoren des Nenners des ursprünglichen Bruches sind. Die einzelnen Brüche nennt man Partialbrüche und die Darstellung eine Partialbruchzerlegung.

Eine solche Partialbruchzerlegung gestaltet sich am einfachsten, wenn man den Nenner $M(x)$ in lauter Primfunktionen zerlegt

$$M(x) = P_1^{e_1}(x)\, P_2^{e_2}(x) \ldots P_n^{e_n}(x)$$

und dann die einzelnen Potenzen der Primfunktionen zu Nennern der Partialbrüche macht, also

$$M_i(x) = P_i^{e_i}(x) \qquad (i = 1, 2, \ldots n)$$

annimmt. In diesem Falle sind die einzelnen Partialbrüche

$$\frac{F_i(x)}{P_i^{e_i}(x)}$$

noch einer Umformung fähig, die wir noch betrachten müssen. Man kann nämlich $F_i(x)$ nach Potenzen von $P_i(x)$ entwickeln. Dividiert man $F_i(x)$ durch $P_i(x)$, so erhält man

$$F_i(x) = F_i^{(1)}(x)\, P_i(x) + F_{i\,e_i - 1}(x),$$

wo $F_{i\,e_i - 1}$ von niederem Grade ist als $P_i(x)$. Man kann so fortfahren und

$$\begin{aligned} F_i^{(1)}(x) &= F_i^{(2)}(x)\, P_i(x) + F_{i\,e_i - 2}(x) \\ F_i^{(2)}(x) &= F_i^{(3)}(x)\, P_i(x) + F_{i\,e_i - 3}(x) \\ &\cdots\cdots \end{aligned}$$

setzen, bis man zuletzt auf

$$F_i^{(e_i - 1)}(x) = F_{i\,0}(x)$$

kommt. Es wird dann

$$\begin{aligned} F_i(x) = F_{i\,0}(x)\, P_i^{e_i - 1}(x) + F_{i\,1}(x)\, P_i^{e_i - 2}(x) + \ldots \\ + F_{i\,e_i - 2}(x)\, P(x) + F_{i\,e_i - 1}(x), \end{aligned}$$

und es kann, da der Grad von $F_i(x)$ kleiner ist als der von

$P_i^{e_i}(x)$, das obige Verfahren nicht mehr fortgesetzt werden. Man erhält nun

$$\frac{F_i(x)}{P_i^{e_i}(x)} = \sum_{k=0}^{e_i-1} \frac{F_{ik}(x)}{P_i^{k+1}(x)}$$

und somit

$$\frac{F(x)}{M(x)} = G(x) + \sum_{i=1}^{i=n} \sum_{k=0}^{k=e_i-1} \frac{F_{ik}(x)}{P_i^{k+1}(x)}.$$

Die Darstellung einer gebrochenen rationalen Funktion in Form von Partialbrüchen erweist sich nützlich bei der Untersuchung ihrer Integration, und wir verweisen, was die Beispiele betrifft, auf Bd. XI dieser Sammlung.

---

§ 67. Anwendung auf die Interpolation.

Bedeutend vereinfachen sich die Entwickelungen des vorigen Paragraphen, wenn man annimmt, dafs die teilerfremden Funktionen $M_i(x)$ alle linear sind. Dieser Fall ist aus dem Grunde der wichtigste, weil, wie sich später zeigen wird (Abschnitt X), sich jede ganze Funktion bei Einführung von Irrationalitäten und komplexen Gröfsen als ein Produkt linearer Faktoren darstellen läfst. Setzen wir nun

$$M_i = x - x_i, \qquad (i = 1, 2, \dots n)$$

so läfst sich jede ganze Funktion $F(x)$ nach M_i auf die Konstante $F(x_i)$ reduzieren (§ 59), und das im vorigen Paragraphen in Nummer 3) behandelte Problem läfst sich daher in anderer Fassung auch so aussprechen: Es soll eine ganze Funktion gefunden werden, die für n gegebene verschiedene Werte der Variabeln $x_1, x_2, \dots x_n$ n vorgeschriebene Werte $y_1 = F(x_1)$, $y_2 = F(x_2), \dots y_n = F(x_n)$ annimmt. Da wir nun wissen, dafs eine solche Funktion nach dem Modul

$$M(x) = (x - x_1)(x - x_2) \dots (x - x_n)$$

völlig bestimmt ist, so ist sie überhaupt völlig bestimmt, wenn ihr Grad kleiner als n sein soll. Wir kommen somit auf dieselbe Interpolationsaufgabe zurück, die wir in § 52 mit Hülfe der Determinanten gelöst hatten, die wir aber jetzt von anderen Gesichtspunkten aus noch einmal behandeln wollen.

Wir wissen, dafs die gesuchte Funktion der Kongruenz

$$X \equiv \sum_i F(x_i) \frac{M(x)}{x - x_i} X_i \mod M(x)$$

$$(i = 1, 2, \ldots n)$$

genügen mufs — da offenbar die im vorigen Paragraphen betrachtete Funktion $L_i(x) = \frac{M(x)}{x - x_i}$ ist — wo X_i durch die Kongruenz

$$L_i(x) X_i \equiv 1 \mod (x - x_i)$$

bestimmt wird. Die Lösung dieser läfst sich aber sofort angeben. Denn da

$$L_i(x) \equiv L_i(x_i) \mod (x - x_i)$$

und ferner $L_i(x_i) = M'(x_i)$ ist, so erhalten wir

$$M'(x_i) X_i \equiv 1 \mod (x - x_i)$$

oder

$$X_i \equiv \frac{1}{M'(x_i)} \mod (x - x_i),$$

so dafs

$$F(x) \equiv \sum_i F(x_i) \frac{M(x)}{(x - x_i) M'(x_i)} \mod M(x)$$

wird. Hieraus folgt aber sofort die Lagrangesche Interpolationsformel in ihrer früheren Gestalt, wenn wir $F(x)$ von niederem Grade als n annehmen. Setzen wir jedoch $F(x)$ von höherem Grade voraus, so können wir die letztere Kongruenz mit Einführung einer ganzen Funktion $G(x)$ als Gleichung darstellen in der Form

$$F(x) = G(x) M(x) + \sum_i F_i(x_i) \frac{M(x)}{(x - x_i) M'(x_i)}$$

und aus ihr die Partialbruchzerlegung

$$\frac{F(x)}{M(x)} = G(x) + \sum_i \frac{F_i(x_i)}{M'(x_i)} \frac{1}{x - x_i}$$

ableiten, die also immer gilt, wenn der Nenner des Bruches $\frac{F(x)}{M(x)}$ in lauter verschiedene Linearfaktoren zerlegbar ist.

VIII. Abschnitt.

Kongruenzen höheren Grades. Quadratische Reste.

---

§. 68. Reduktion der Kongruenzen hinsichtlich des Moduls.

Bezeichnen wir mit $F(x)$ eine ganze ganzzahlige Funktion von x, so hat eine Kongruenz höheren Grades von der Form

$$F(x) \equiv 0 \mod m$$

immer unendlich viele ganze Zahlen x als Lösung, wenn sie überhaupt Lösungen zuläfst; denn wenn x_0 eine Lösung ist, so sind alle Zahlen, die der Bedingung

$$x \equiv x_0 \mod m$$

genügen und in unendlicher Anzahl vorhanden sind, ebenfalls solche, weil immer

$$F(x) \equiv F(x_0) \mod m$$

ist. Das Problem der Auflösung der Kongruenzen kommt also darauf hinaus, die sämtlichen inkongruenten Lösungssysteme nach dem Modul m zu finden, wie wir das schon früher in § 29 bemerkt haben bei der Behandlung der linearen Kongruenzen.

Das Problem läfst sich nun bedeutend vereinfachen, was den Modul m betrifft. Wir nehmen an, dafs m als ein Produkt

$$m = m_1\, m_2 \ldots m_n$$

von lauter Zahlen dargestellt ist, von denen immer zwei gegen einander relativ prim sind, wie es z. B. der Fall ist, wenn wir für $m_1, m_2, \dots m_n$ die sämtlichen in m enthaltenen Primzahlpotenzen wählen. Aus $F(x) \equiv 0 \bmod m$ folgt nun sofort

$$F(x) \equiv 0 \bmod m_i. \qquad (i = 1, 2, \dots n)$$

Ist also eine dieser Kongruenzen unlösbar, so ist es auch die ursprüngliche. Aber auch das Umgekehrte gilt, wie wir jetzt beweisen wollen. Haben die Kongruenzen

$$F(x_i) \equiv 0 \bmod m_i \qquad (i = 1, 2, \dots n)$$

allgemein ein Lösungssystem

$$x_i \equiv a_i \bmod m_i, \qquad (i = 1, 2, \dots n)$$

und bestimmt man dann x so, dafs

$$x \equiv x_i \equiv a_i \bmod m_i \qquad (i = 1, 2, \dots n)$$

wird (§ 31), so folgt aus

$$F(x) \equiv F(x_i) \equiv 0 \bmod m_i \qquad (i = 1, 2, \dots n)$$

sofort

$$F(x) \equiv 0 \bmod m = m_1 \, m_2 \dots m_n,$$

wie zu zeigen war. Auch bezüglich der Anzahl der Wurzeln dieser Kongruenz gelangen wir zu einem einfachen Resultate. Hat nämlich die Kongruenz

$$F(x_i) \equiv 0 \bmod m_i \qquad (i = 1, 2, \dots n)$$

λ_i inkongruente Lösungssysteme

$$x_i \equiv a_{i h_i} \bmod m_i, \qquad \begin{matrix}(i = 1, 2, \dots n; \\ h_i = 1, 2, \dots \lambda_i)\end{matrix}$$

so hat x, weil jedem System $x_1, x_2, \dots x_n$ ein einziger Wert mod m entspricht und zwei verschiedenen auch inkongruente mod m, im Ganzen $\lambda_1 \, \lambda_2 \dots \lambda_n$ Werte nach dem Modul m.

Als Resultat der Untersuchung ergiebt sich daher, dafs wir nur Kongruenzen nach Moduln zu betrachten brauchen, die Potenzen von Primzahlen sind. Wenn aber eine Kongruenz nach einer Primzahlpotenz als Modul gilt, so mufs sie erst recht für die Primzahl selbst Gültigkeit haben, so dafs wir genötigt sind, uns zunächst auf Kongruenzen nach einem Primzahlmodul zu beschränken.

§ 69. Modulsysteme von ganzen ganzzahligen Funktionen mit einer Primzahl.

Eine jede Kongruenz läfst sich immer als eine Äquivalenz schreiben, so dafs

$$F(x) \equiv 0 \mod m$$

auch durch

$$[F(x),\ m] = m$$

ersetzt werden kann. Wenn der Modul m gleich einer Primzahl p ist, so läfst sich das auf der linken Seite auftretende Modulsystem

$$[F(x),\ p]$$

vereinfachen. Zunächst kann man aus der ganzen ganzzahligen Funktion

$$F(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

alle durch p teilbaren Glieder weglassen und dann die Funktion so umformen, dafs der Koeffizient der höchsten Potenz der Einheit gleich wird. Nehmen wir an, dafs a_0 nicht durch p teilbar, also $(a_0, p) = 1$ sei, so kann man eine Zahl $\overline{a_0}$ so bestimmen, dafs

$$a_0 \overline{a_0} \equiv 1 \mod p$$

ist. Es ist dann aber

$$[F(x),\ p] = [(\overline{a_0},\ p)\ F(x),\ p] = [\overline{a_0}\, F(x),\ p],$$

und $\overline{a_0}\, F(x)$ hat, wenn man $a_0 \overline{a_0}$ nach p reduziert, die gewünschte Form. Übrigens hätte man etwas umständlicher dasselbe ableiten können, wenn man dem Modulsystem das Element $p x^n$ hinzugefügt und dann das Teilsystem $[F(x), p x^n]$ reduziert hätte. In dem betrachteten Modulsystem tritt nur eine einzige ganze ganzzahlige Funktion der unbestimmten Gröfse x auf. Es läfst sich nun zeigen, dafs ein Modulsystem mit mehreren solchen Funktionen

$$[A(x),\ B(x),\ C(x), \ldots, p]$$

stets auf die Form

$$[T(x),\ p]$$

gebracht werden kann, in der nur eine einzige Funktion vorkommt.

Sind nämlich A(x) und B(x) als ganze ganzzahlige Funktionen so vorbereitet, dafs in ihnen die Koeffizienten der höchsten Potenz der Unbestimmten der Einheit gleich sind, sind diese Potenzen die α te in A(x), die β te in B(x), so kann man, wenn $\alpha \geqq \beta$ angenommen wird, die Zerlegung

$$A(x) = x^{\alpha-\beta} B(x) + A_1(x)$$

ausführen, so dafs der Grad von $A_1(x)$ kleiner als α wird. (Vgl. § 56.) Dann wird

$$[A(x), B(x), C(x), \ldots, p] = [A_1(x), B(x), C(x), \ldots, p],$$

und man kann nun aus $A_1(x)$ alle etwa durch p teilbaren Koeffizienten fortschaffen, wobei sich der Grad unter Umständen noch weiter erniedrigt, und es dann so umformen, dafs der Koeffizient der höchsten Potenz wieder der Einheit gleich wird. Mit dem so umgewandelten System kann man aber genau so verfahren wie mit dem ursprünglichen. Stets wird der Grad von einer der Funktionen erniedrigt, und deswegen mufs man nach einer beschränkten Anzahl von Reduktionen schliefslich zu einem Modulsystem gelangen, in dem nur eine einzige Funktion T(x) vorkommt. Es wird dann

$$[A(x), B(x), C(x), \ldots, p] = [T(x), p].$$

T(x) kann auch eine ganze Zahl sein, und in diesem Falle wird das Modulsystem entweder äquivalent p oder der Einheit, je nachdem die ganze Zahl durch p teilbar ist oder nicht.

Sind a und b inkongruent nach p, so ist

$$(x - a, x - b, p) = (x - a, a - b, p) = 1.$$

---

§ 70. Anzahl der Wurzeln einer Kongruenz nach einem Primzahlmodul.

Da F(x) — F(a) stets durch x — a teilbar ist, so ist das Modulsystem

$$[F(x), x - a, p] = [F(a), x - a, p]$$

äquivalent (x — a, p) oder 1, je nachdem die Kongruenz

$$F(x) \equiv 0 \mod p$$

die Lösung

$$x \equiv a \mod p$$

besitzt oder nicht. Wir wollen nun voraussetzen, dafs die Kongruenz eine Reihe von m inkongruenten Lösungssystemen

$$x \equiv a_i \bmod p \qquad (i = 1, 2, \ldots m)$$

besitzt, m darf hierbei den Wert p nicht überschreiten, und dann untersuchen, in welcher Beziehung F(x) zu den Linearfaktoren $x - a_1, x - a_2, \ldots x - a_m$ steht. Dazu schicken wir folgenden Satz vorauf:

Ist

$$[A(x), B(x), p] = 1,$$

so ist immer

$$[A(x)\,B(x), C(x), p] = [A(x), C(x), p]\,[B(x), C(x), p].$$

In der That läfst sich das Modulsystem auf der rechten Seite schreiben in der Form

$$[A(x)B(x), A(x)C(x), B(x)C(x), A(x)p, B(x)p, p^2, C(x)p, C^2(x)].$$

Bedenkt man nun aber, dafs

$$[A(x)\,C(x), B(x)\,C(x), C(x)\,p] = [A(x), B(x), p]\,C(x) = C(x)$$

und

$$[A(x)\,p, B(x)\,p, p^2] = [A(x), B(x), p]\,p = p$$

ist, so erhält man

$$[A(x)\,B(x), C(x), C^2(x), p]$$

und daraus nach Fortlassung von $C^2(x)$ die linke Seite der angegebenen Äquivalenz.

Ebenso wie in § 20 läfst sich nun der allgemeinere Satz ableiten:

Wenn $A_1(x), A_2(x), \ldots A_m(x)$ die Bedingung erfüllen, dafs für $i \neq k$ stets

$$[A_i(x), A_k(x), p] = 1$$

ist, so ist

$$[A_1(x)\,A_2(x) \ldots A_m(x), B(x), p]$$
$$= [A_1(x), B(x), p]\,[A_2(x), B(x), p] \ldots [A_m(x), B(x), p].$$

Wenden wir nun diesen Satz auf den oben angegebenen Fall an, indem wir

$$A_i(x) = x - a_i$$
$$B(x) = F(x) \text{ und } B(x) = p$$

annehmen, so finden wir

$$\begin{aligned}
&[F(x), (x - a_1)(x - a_2)\dots(x - a_m, p] \\
&= [F(x), x - a_1, p]\,[F(x), x - a_2, p]\dots[F(x), x - a_m, p] \\
&= (x - a_1, p)(x - a_2, p)\dots(x - a_m, p) \\
&[(x - a_1)(x - a_2)\dots(x - a_m), p] \\
&= (x - a_1, p)(x - a_2, p)\dots(x - a_m, p),
\end{aligned}$$

so dafs

$$\begin{aligned}
&[F(x), (x - a_1)(x - a_2)\dots(x - a_m), p] \\
&= [(x - a_1)(x - a_2)\dots(x - a_m), p]
\end{aligned}$$

wird. Diese Äquivalenz besagt aber, dafs sich $F(x)$ in der Form

$$F(x) = (x - a_1)(x - a_2)\dots(x - a_m)\,\varphi(x) + p\,\psi(x)$$

darstellen lassen mufs, wo $\varphi(x)$ und $\psi(x)$ zwei ganze ganzzahlige Funktionen von x darstellen, und daraus folgt, dafs der Grad n von $F(x)$ nicht kleiner sein kann als die Anzahl m der inkongruenten Lösungssysteme.

Wir machen von diesem Satze sofort eine Anwendung auf die Kongruenz

$$x^p \equiv x \mod p,$$

der zufolge des Fermatschen Satzes alle inkongruenten Zahlen genügen. Es ergiebt sich dann sofort, dafs sich $x^p - x$, abgesehen von einer mit p multiplizierten ganzen ganzzahlige Funktion, in der Form

$$x(x - 1)(x - 2)\dots(x - p + 1)$$

darstellen lassen mufs, oder dafs

$$x^p - x \equiv x(x - 1)(x - 2)\dots(x - p + 1) \mod p$$

ist. Hierbei müssen die Koeffizienten der gleichen Potenzen auf beiden Seiten nach dem Modul p kongruent sein. Der Koeffizient von der ersten Potenz auf der linken Seite ist -1, auf der rechten $(-1)^{p-1}\,1 . 2 . 3 \dots (p - 1)$; es ist daher, wenn $p \neq 2$ ist,

$$1 . 2 . 3 \dots (p - 1) + 1 \equiv 0 \mod p;$$

aber es ist leicht einzusehen, dafs diese Kongruenz auch für den Fall $p = 2$ noch gültig ist, weil $+1 \equiv -1 \mod 2$ ist. Die Kongruenz drückt den Inhalt des Wilsonschen Satzes aus:

Das um die Einheit vermehrte Produkt aller zu einer Primzahl teilerfremden inkongruenten Reste ist durch die Primzahl teilbar.

Dieser Satz ist auch umkehrbar: Wenn

$$1 . 2 . 3 \ldots (p-1) + 1 \equiv 0 \mod p$$

ist, so mufs p eine Primzahl sein. Wäre nämlich p zusammengesetzt und durch q teilbar, das mit einer der Zahlen $2, 3, \ldots (p-1)$ übereinstimmt, so kann die obige Summe nicht durch q teilbar sein, weil 1 nicht dadurch teilbar ist, und also auch nicht durch p.

Wenn es sich darum handelt, eine Kongruenz

$$F(x) \equiv 0 \mod p$$

zu lösen, so kann man dem Modulsystem $[F(x), p]$ stets die Funktion $x^p - x$ beifügen. Das so erhaltene System

$$[F(x), x^p - x, p]$$

läfst sich aber, wie wir in § 69 gesehen haben, stets auf die Form

$$[\Phi(x), p]$$

reduzieren, wo $\Phi(x)$ eine ganze ganzzahlige Funktion ist, die in $x^p - x$ enthalten ist, wenn man von Vielfachen von p absieht, und also die Form hat

$$\Phi(x) \equiv (x - a_1)(x - a_2) \ldots (x - a_m) \mod p,$$

wo $a_1, a_2, \ldots a_m$ nach p inkongruente Reste bedeuten. Alle Wurzeln der Kongruenz $F(x) \equiv 0 \mod p$ stimmen demnach völlig überein mit denen der Kongruenz $\Phi(x) \equiv 0 \mod p$. Hiermit ist die Frage nach der Anzahl der Lösungssysteme vollständig beantwortet.

---

§ 71. Gruppen im verkürzten Restsystem nach einem Primzahlmodul.

Besonders einfache Folgerungen lassen sich ziehen, wenn $F(x)$ die Form $x^m - 1$ hat. Es ist allgemein

$$(x^a - 1, x^b - 1) = x^{(a, b)} - 1,$$

und hieraus ergeben sich folgende Resultate:

1) Setzen wir $a = m$, $b = p - 1$, so haben wir den Satz: Sämtliche Wurzeln der Kongruenz

$$x^m \equiv 1 \mod p$$

genügen auch der Kongruenz

$$x^{(m, p-1)} \equiv 1 \mod p,$$

die soviele inkongruenten Wurzeln hat als ihr Grad anzeigt.

2) Ist also m ein Teiler von $p - 1$, so hat die Kongruenz

$$x^m \equiv 1 \mod p$$

immer m inkongruente Wurzeln.

3) Behalten wir die Voraussetzung über m bei und setzen $a = m$, $b = rm$, so dafs $(a, b) = m$ wird, so folgt, dafs man durch Potenzierung jeder Wurzel der Kongruenz

$$x^m \equiv 1 \mod p \qquad ((m, p-1) = m)$$

wieder eine Wurzel derselben Kongruenz erhält.

4) Wir haben nun früher bei der Betrachtung des verkürzten Restsystems gesehen (§ 26), dafs es zu jeder Zahl r in ihr einen kleinsten Exponenten m giebt, so dafs

$$r^m \equiv 1 \mod p, \qquad ((m, p-1) = m)$$

ist, und die m inkongruenten Zahlen

$$1, r, r^2, \ldots r^{m-1}$$

eine Gruppe darstellen. Alle Zahlen dieser Gruppe sind nach 3) Wurzeln der Kongruenz

$$x^m \equiv 1 \mod p,$$

die nach 2) keine andern Wurzeln mehr besitzt. Wenn wir uns nun zunächst die Frage vorlegen, ob es noch andere Grundelemente als r giebt, durch deren Potenzierung die ganze Gruppe erzeugt werden kann, so ist demnach klar, dafs als solche nur Potenzen von r in Betracht kommen. Betrachten wir nun irgend eine Potenz r^h und nehmen an, es wäre der zu r^h gehörige Exponent gleich n, so folgt aus unserer obigen Formel, wenn wir $a = m$, $b = rn$ setzen, dafs r der Kongruenz

$$x^{(m, hn)} \equiv 1 \mod p$$

genügen mufs. Da aber nach unserer Annahme m der kleinste Exponent ist, für den

$$r^m \equiv 1 \mod p$$

ist, so folgt

$$(h n, m) = m$$

$$\left[n, \frac{m}{(h, m)}\right] = \frac{m}{(h, m)}.$$

Der Exponent n, der zu r^h gehört, ist also ein Vielfaches von $\frac{m}{(m, h)}$. Nun ist aber

$$(r^h)^{\frac{m}{(m, h)}} = r^{\frac{m h}{(m, h)}}$$

und $\frac{m h}{(m, h)}$ ein Vielfaches von m, daraus folgt

$$(r^h)^{\frac{m}{(m, h)}} \equiv 1 \mod p$$

und somit, dafs $n = \frac{m}{(m, h)}$ ist.

Soll r^h zum Exponenten m gehören, so mufs $(h, m) = 1$ sein, so dafs h also $\varphi(m)$ inkongruente Werte nach dem Modul m annehmen kann. Wir wollen nun alle Wurzeln der Kongruenz

$$x^m \equiv 1 \mod p, \qquad ((m, p-1) = m)$$

die keiner Kongruenz von derselben Form, aber niederem Grade

$$x^n \equiv 1 \mod p \qquad (n < m)$$

genügen können, ihre primitive Wurzel nennen. Dann können wir das Resultat unserer Untersuchung so formulieren:

Wenn die Kongruenz

$$x^m \equiv 1 \mod p \qquad ((m, p-1) = m)$$

überhaupt primitive Wurzeln besitzt, so hat sie $\varphi(m)$ inkongruente primitive Wurzeln und nicht mehr.

5) Unsere Untersuchung ist nun aber noch insofern unvollständig, als daraus nicht hervorgeht, ob eine Kongruenz von obiger Form überhaupt primitive Wurzeln hat oder nicht. Um diese Frage zu entscheiden, betrachten wir das ganze verkürzte Restsystem nach dem Modul p und teilen die

Reste in Abteilungen ein, so dafs alle in derselben Abteilung denselben Exponenten d haben, der ein Teiler von (p — 1) sein mufs. Da in der Abteilung mit dem Exponenten d nun $\varphi(d)$ Reste vorkommen, so besteht das verkürzte Restsystem aus

$$\sum_{(d)} \varphi(d)$$

Resten, wobei die Summation auf alle solche Teiler d von p — 1 zu erstrecken ist, deren zugehörige Kongruenzen

$$x^d \equiv 1 \bmod p$$

überhaupt primitive Wurzeln haben. Wenn das nun für einige Teiler von d nicht zuträfe, so würde das verkürzte Restsystem weniger als (p — 1) Elemente besitzen, da sich die Summe erst dann gleich (p — 1) ergiebt, wenn die Summation auf alle Teiler d von (p — 1) bezieht. Daher erhalten wir den Satz:

Ist $(m, p-1) = m$, so hat die Kongruenz

$$x^m \equiv 1 \bmod p$$

immer $\varphi(m)$ primitive Wurzeln.

---

§ 73. Primitive Wurzeln. Indizes.

Setzen wir $m = p - 1$, so erhalten wir aus dem letzten Satz als spezielle Folgerung:

Die Kongruenz

$$x^{p-1} \equiv 1 \bmod p$$

hat stets $\varphi(p-1)$ primitive Wurzeln. Diese nennt man auch häufig kurz primitive Wurzeln der Primzahl p.

Ist g eine solche, so sind die Potenzen

$$1, g, g^2, \dots g^{p-2}$$

sämtlich inkongruent und stellen daher die ganze verkürzte Restgruppe nach dem Primzahlmodul p dar. Ist also a eine durch p nicht teilbare Zahl, so kann man immer eine Zahl $\alpha \bmod p - 1$ bestimmen, so dafs

$$g^\alpha \equiv a \bmod p$$

ist. Ist z. B. $p = 19$, so ist 2 eine primitive Wurzel, und

man erhält folgende Tafel, in der die erste Reihe die Exponenten der primitiven Wurzel, die zweite die entsprechenden Potenzreste enthält:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	2	4	8	16	13	7	14	9	18	17	15	11	3	6	12	5	10

Wenn man in dieser Weise das verkürzte Restsystem durch die Potenzen eines Grundelementes g darstellt, so nennt man den Exponenten α auch den *Index von a in Bezug auf die Basis g* und schreibt

$$\alpha \equiv {}^{g}ind\, a \mod p-1,$$

so dafs

$$g^{{}^{g}ind\, a} \equiv a \mod p$$

ist. Speziell ist also für jede Basis $ind\, 1 \equiv 0 \mod p-1$. Aus

$$g^{\alpha} \equiv a \mod p, \quad g^{\beta} \equiv b \mod p$$

folgt

$$g^{\alpha+\beta} \equiv g^{\alpha} g^{\beta} \equiv a\, b,$$

oder anders ausgedrückt: Es ist

$$ind\; a\, b \equiv ind\; a + ind\; b \mod p-1,$$

wobei die Basis g, da sie als unveränderlich vorausgesetzt wird, in der Bezeichnung weggelassen ist.

Der Index eines Produkts ist also kongruent der Summe der Indizes der einzelnen Faktoren.

Dieser Satz läfst sich leicht auf ein Produkt von mehreren Faktoren verallgemeinern, so dafs man allgemein hat

$$ind\; a\, b\, c \ldots \equiv ind\; a + ind\; b + ind\; c + \ldots \mod p-1.$$

Nimmt man an, dafs alle Faktoren gleich a und in der Anzahl n vorhanden sind, so ergiebt sich speziell

$$ind\; a^{n} \equiv n\; ind\; a \mod p-1,$$

oder *man erhält den Index einer Potenz durch Multiplikation des Exponenten mit dem Index der Basis.*

Da $(-1)^2 = 1$ ist, so ist also für jede Basis $2\, ind\, (-1) \equiv 0 \mod p-1$, und daraus folgt $ind\,(-1) \equiv 0 \mod \frac{p-1}{2}$. Da aber $ind\,(-1)$ nicht $\equiv 0 \mod p-1$ sein kann, so folgt allgemein $ind\,(-1) \equiv \frac{p-1}{2} \mod p-1$.

Die abgeleiteten Gesetze ermöglichen es, den Index jeder beliebigen Zahl aus den Indizes aller in ihr enthaltenen Primzahlen zu berechnen. Will man also eine Tafel herstellen, aus der man alle Indizes ableiten kann, so braucht man nur alle Primzahlen zu berücksichtigen, die kleiner als p sind. Für $p = 19$ genügt unter Zugrundelegung von 2 als primitiver Wurzel die folgende Tafel, die aus der oben hingeschriebenen leicht konstruiert werden kann:

Zahl	1	2	3	5	7	11	13	17
Index	0	1	13	16	6	12	5	10

Es ist z. B.

$$ind\,12 \equiv ind\,2^2 + ind\,3 \equiv 2\,ind\,2 + ind\,3 \equiv 15 \bmod 18$$
$$ind\,16 \equiv 4\,ind\,2 \equiv 4 \bmod 18$$
$$ind\,18 \equiv ind\,2 + 2\,ind\,3 \equiv 9 \bmod 18.$$

Auf der folgenden Seite findet sich eine solche Tafel beigefügt, deren Gebrauch keiner weiteren Erläuterung bedarf. Als primitive Wurzeln sind dabei nicht überall die kleinsten zu Grunde gelegt; mit Rücksicht auf gewisse Anwendungen ist es nämlich vorteilhaft, die Grundzahl 10 unseres Zahlsystems so zu bevorzugen, dafs sie einen möglichst kleinen Index erhält, und daher, wenn sie primitive Wurzel ist, als Basis zu Grunde zu legen. Was die kleinsten primitiven Wurzeln anbelangt, so bemerken wir, dafs diese

gleich 2 ist für $p = 3, 5, 11, 13, 19, 29, 37, 53, 59, 61, 67, 83$
3 . . . $p = 7, 17, 31, 43, 79, 89$
5 . . . $p = 23, 47, 73, 97$
6 . . . $p = 41$
7 . . . $p = 71.$

Mit Hülfe der primitiven Wurzeln ist es leicht, alle im verkürzten Restsystem enthaltenen Gruppen aufzustellen. Die folgenden Beispiele werden leicht verständlich sein; die Grundelemente sind durch fetten Druck hervorgehoben.

$$\mathbf{p = 13}$$
$$G_6 = (1,\ 3,\ \mathbf{4},\ 9,\ \mathbf{10},\ 12)$$
$$G_4 = (1,\ \mathbf{5},\ \mathbf{8},\ 12)$$
$$G_3 = (1,\ \mathbf{3},\ \mathbf{9})$$
$$G_2 = (1,\ \mathbf{12}).$$

Indizes-Tafel nach Gaufs.

Modul	Basis	Numeri 2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	59	61	67	71	73	79	83	89
		Indizes																							
3	2	1																							
5	2	1	3																						
7	3	2	1	5																					
11	2	1	8	4	7																				
13	6	5	8	9	7	11																			
17	10	10	11	7	9	13	12																		
19	10	17	5	2	12	6	13	8																	
23	10	8	20	15	21	3	12	17	5																
29	10	11	27	18	20	23	2	7	15	24															
31	17	12	13	20	4	29	23	1	22	21	27														
37	5	11	34	1	28	6	13	5	25	21	15	27													
41	6	26	15	22	39	3	31	33	9	36	7	28	32												
43	28	39	17	5	7	6	40	16	29	20	25	32	35	18											
47	10	30	18	17	38	27	3	42	29	39	43	5	24	25	37										
53	26	25	9	31	38	46	28	42	41	39	6	45	22	33	30	8									
59	10	25	32	34	44	45	23	14	22	27	4	7	41	2	13	53	28								
61	10	47	42	14	23	45	20	49	22	39	25	13	33	18	41	40	51	17							
67	12	29	9	39	7	61	23	8	26	20	22	43	44	19	63	64	3	54	5						
71	62	58	18	14	33	43	27	7	38	5	4	13	30	55	44	17	59	29	37	11					
73	5	8	6	1	33	55	59	21	62	46	35	11	64	4	51	31	53	5	58	50	44				
79	29	50	71	34	19	70	74	9	10	52	1	76	23	21	47	55	7	17	75	54	33	4			
83	50	3	52	81	24	72	67	4	59	16	36	32	60	38	49	69	13	20	34	53	17	43	47		
89	30	72	87	18	7	4	65	82	53	31	29	57	77	67	59	34	10	45	19	32	26	68	46	27	
97	10	86	2	11	53	82	83	19	27	79	47	26	41	71	44	60	14	65	32	51	25	20	42	91	18

$p = 17$
$G_8 = (1, 2, 4, 8, 9, 13, 15, 16)$
$G_4 = (1, 4, 13, 16)$
$G_2 = (1, 16).$

$p = 19$
$G_9 = (1, 4, 5, 6, 7, 9, 11, 16, 17)$
$G_6 = (1, 7, 8, 11, 12, 18)$
$G_3 = (1, 7, 11)$
$G_2 = (1, 18).$

---

§ 73. Binomische Kongruenzen.

Wir wenden uns jetzt zu einer Verallgemeinerung der bisherigen Kongruenzen und betrachten solche von der Form

$$x^n \equiv a \bmod p,$$

die man auch binomische oder reine Kongruenzen nennt. Hierbei soll a als teilerfremd zu p vorausgesetzt werden; ist das nicht der Fall, so lautet die einzige mögliche Lösung $x \equiv 0 \bmod p$. Um die Möglichkeit der Lösung und die Anzahl der Lösungssysteme zu beurteilen, haben wir nach § 70 das Modulsystem

$$(x^n - a,\ x^{p-1} - 1)$$

genauer zu untersuchen. Wir wollen dies ganz allgemein durchführen und an seiner Stelle das System

$$(x^a - 1,\ x^b - y)$$

betrachten. Hier ist zunächst klar, daſs dem Modulsystem die Funktion $x^{\frac{ab}{(a,b)}} - 1$ hinzugefügt werden kann, die durch $x^a - 1$ teilbar ist. Da aber

$$x^{\frac{ab}{(a,b)}} - 1 = \left(x^{\frac{ab}{(a,b)}} - y^{\frac{a}{(a,b)}}\right) + \left(y^{\frac{a}{(a,b)}} - 1\right)$$

ist, und das erste Glied auf der rechten Seite durch $x^b - y$ teilbar ist, so ist unser Modulsystem auch dem folgenden äquivalent

$$[x^a - 1,\ x^b - y,\ y^{\frac{a}{(a,b)}} - 1].$$

Die beiden ersten Elemente formen wir nun noch etwas um. Wir denken uns eine Gröfse h bestimmt, die der Kongruenz

$$\frac{b}{(a, b)} h \equiv 1 \bmod \frac{a}{(a, b)}$$

genügt.

Dann können wir dem Modulsystem als Element

$$x^{(a, b)} - y^h = (x^{bh} - y^h) - (x^{bh} - x^{(a, b)})$$

hinzufügen, da $x^{bh} - y^h$ durch $x^b - y$, $x^{bh} - x^{(a, b)} = x^{(a, b)} \left[x^{\left(\frac{bh}{(a, b)} - 1\right)(a, b)} - 1\right]$ durch $x^a - 1$ teilbar ist, und dann an Stelle von $x^a - 1$ und $x^b - y$ bezw. $x^a - y^{\frac{ah}{(a, b)}}$ $x^b - y^{\frac{bh}{(a, b)}}$ setzen, denn es ist

$$x^a - 1 = \left(x^a - y^{\frac{ah}{(a, b)}}\right) + \left(y^{\frac{ah}{(a, b)}} - 1\right)$$

$$x^b - y = \left(x^b - y^{\frac{bh}{(a, b)}}\right) + \left(y^{\frac{hb}{(a, b)} - 1} - 1\right) y,$$

und die beiden letzten Glieder auf den rechten Seiten sind durch $y^{\frac{a}{(a, b)}} - 1$ teilbar. Unser Modulsystem geht sonach in das folgende über

$$\left[x^{(a, b)} - y^h, \ x^a - y^{\frac{ah}{(a, b)}}, \ x^b - y^{\frac{bh}{(a, b)}}, \ y^{\frac{a}{(a, b)}} - 1\right].$$

Da nun aber das zweite und das dritte Element durch das erste teilbar sind, so können sie entfernt werden, und wir erhalten schliefslich die Äquivalenz

$$(x^a - 1, \ x^b - y) = \left[x^{(a, b)} - y^h, \ y^{\frac{a}{(a, b)}} - 1\right].$$

Wenden wir nun das erhaltene Resultat auf das zu Anfang erwähnte Modulsystem an, so finden wir

$$(x^n - a, \ x^{p-1} - 1) = \left[x^{(p-1, n)} - a^h, \ a^{\frac{p-1}{(p-1, n)}} - 1\right],$$

wo h der Kongruenz

$$\frac{nh}{(p-1, n)} \equiv 1 \bmod \frac{p-1}{(p-1, n)}$$

genügt. Hieraus ergiebt sich sofort der Satz (§ 70, Schlufssatz):

Die hinreichende und notwendige Bedingung für die Lösbarkeit der Kongruenz

$$x^n \equiv a \mod p$$

lautet

$$a^{\frac{p-1}{(p-1,n)}} \equiv 1 \mod p.$$

Ist diese erfüllt, so hat sie $(p-1, n)$ inkongruente Lösungssysteme nach dem Modul p.

Ist die Kongruenz möglich, so nennt man a einen nten Potenzrest nach der Primzahl p. Da ein solcher der Kongruenz

$$x^{\frac{p-1}{(p-1,n)}} \equiv 1 \mod p$$

genügt, die nach § 71, 2 ebensoviele inkongruente Lösungssysteme hat, wie ihr Grad beträgt, so giebt es $\frac{p-1}{(p-1, n)}$ verschiedene inkongruente nte Potenzreste der Primzahl p.

Man kann die Kongruenz

$$x^n \equiv a \mod p$$

auch mit Hülfe der primitiven Wurzeln, oder was auf dasselbe hinauskommt, mit Hülfe der Indizes behandeln. Setzt man nämlich, wenn g eine primitive Wurzel von p bedeutet,

$$x \equiv g^{\xi}, \qquad a \equiv g^{\alpha} \mod p$$

oder

$$\xi \equiv \mathit{ind}\, x, \quad \alpha \equiv \mathit{ind}\, a \mod p-1,$$

so folgt aus ihr

$$n\,\xi \equiv \alpha \mod p-1.$$

Die hinreichende und notwendige Bedingung für die Lösbarkeit lautet daher nach § 29

$$(\alpha, n, p-1) = (n, p-1),$$

und wenn sie erfüllt ist, so giebt es für ξ $(p-1, n)$ Lösungssysteme nach dem Modul $p-1$; da aber jedem ein bestimmter Wert für x nach p als Modul entspricht, so hat die Kongruenz $x^n \equiv a$ auch $(p-1, n)$ Lösungssysteme nach dem Modul p. Man sieht auch ferner leicht ein, dafs die Bedingung

$$(\alpha, p-1, n) = (p-1, n)$$

völlig gleichbedeutend ist mit

$$a^{\frac{p-1}{(p-1,n)}} \equiv 1 \mod p,$$

denn ist α durch $(p-1, n)$ teilbar, so ist

$$a^{\frac{p-1}{(p-1,n)}} \equiv g^{\frac{\alpha(p-1)}{(p-1,n)}} \equiv 1 \mod p$$

und umgekehrt, ist $a = g^\alpha$ und $a^{\frac{p-1}{(p-1,n)}} \equiv 1 \mod p$, so ist

$$g^{\frac{\alpha(p-1)}{(p-1,n)}} \equiv 1 \mod p$$

nur möglich, wenn der Exponent $\frac{\alpha(p-1)}{(p-1,n)}$ durch $p-1$, also α durch $(p-1, n)$ teilbar ist.

Ein Beispiel möge die Benutzung der Indizes veranschaulichen:

$$x^{12} \equiv 7 \mod 19.$$

Nehmen wir die primitive Wurzel 2 als Basis und setzen *ind* $x = \xi$, so erhalten wir

$$\begin{aligned} 12\,\xi &\equiv 6 \mod 18 \\ 2\,\xi &\equiv 1 \mod 3 \\ \xi &\equiv 2 \mod 3 \\ \xi &\equiv 2, 5, 8, 11, 14, 17 \mod 18 \end{aligned}$$

und daraus

$$x \equiv 4, 13, 9, 15, 6, 10 \mod 19.$$

---

§ 74. Quadratische Reste und Nichtreste.

Einen zweiten Potenzrest nennt man auch einen **quadratischen Rest**. Da es $\frac{p-1}{2}$ quadratische Reste giebt, die der Kongruenz

$$x^{\frac{p-1}{2}} \equiv 1 \mod p$$

genügen, so genügen alle übrigen $\frac{p-1}{2}$ Elemente des ver-

kürzten Restsystems, die man auch als quadratische Nichtreste bezeichnet, der Kongruenz

$$x^{\frac{p-1}{2}} \equiv -1 \bmod p,$$

wie sich einfach daraus ergiebt, dafs

$$x^{p-1} - 1 = \left(x^{\frac{p-1}{2}} - 1\right)\left(x^{\frac{p-1}{2}} + 1\right)$$

immer durch p teilbar ist, so dafs, wenn von den beiden komplementären Teilern der eine nicht durch p teilbar ist, der andre es sein mufs.

Alle quadratischen Reste bilden, wie überhaupt die Potenzreste, eine in dem verkürzten Restsystem enthaltene Gruppe, so dafs das Produkt von quadratischen Resten wieder ein solcher ist. Beachtet man aber, dafs

$$(a\,b\,c\ldots)^{\frac{p-1}{2}} = a^{\frac{p-1}{2}}\, b^{\frac{p-1}{2}}\, c^{\frac{p-1}{2}} \ldots,$$

und also $(a\,b\,c\ldots)^{\frac{p-1}{2}}$ kongruent $+1$ oder -1 mod p ist, je nachdem die Faktoren $a^{\frac{p-1}{2}}$, $b^{\frac{p-1}{2}}$, $c^{\frac{p-1}{2}}$ ..., die kongruent -1 sind, in einer geraden oder ungeraden Anzahl vorhanden sind, so erhalten wir als Verallgemeinerung den Satz:

Das Produkt aus mehreren Zahlen des verkürzten Restsystems ist ein quadratischer Rest oder Nichtrest, je nachdem unter den Faktoren eine gerade oder ungerade Anzahl von Nichtresten vorkommen.

Speziell ist also das Produkt aus einem Rest und einem Nichtrest ein Nichtrest, während das Produkt zweier Nichtreste wieder ein Rest wird. Man kann übrigens diese speziellen Folgerungen aus der Eigenschaft der quadratischen Reste, eine Gruppe von der Ordnung $\frac{p-1}{2}$ zu bilden, ableiten (wie in § 26) und dann leicht zu dem eben angegebenen allgemeinen Satze aufsteigen.

Die sämtlichen quadratischen Reste einer Primzahl kann man sofort hinschreiben, indem man die $\frac{p-1}{2}$ Quadrate

$$1^2,\ 2^2, \ldots \left(\frac{p-1}{2}\right)^2$$

nach dem Modul p reduziert; die folgenden Quadrate geben nämlich genau dieselben Reste, weil allgemein $(p - r)^2 \equiv r^2$ mod p ist. Übrigens kann man auch sonst leicht nachweisen, daſs die Quadrate auch lauter verschiedene Reste liefern.

Bedeutet g eine primitive Wurzel von p, so liefert die Reihe ihrer Potenzen mit geraden Exponenten

$$g^0 = 1,\ g^2,\ g^4, \ldots g^{p-3}$$

die sämtlichen quadratischen Reste, während in der Reihe mit ungeraden Exponenten

$$g,\ g^3,\ g^5, \ldots g^{p-2}$$

die sämtlichen Nichtreste zur Darstellung kommen.

Ist a ein quadratischer Rest, so lassen sich die Wurzeln der Kongruenz

$$x^2 \equiv a \mod p$$

mit Hülfe der Indizes leicht finden, wie wir im vorigen Paragraphen dargelegt haben. Ist übrigens eine Wurzel $x \equiv x_0$, so muſs die andere $x \equiv - x_0$ und p sein, weil $(- x_0)^2 \equiv {x_0}^2$ mod p ist. Wenn p eine Primzahl ist, die der Kongruenz

$$p \equiv 3 \mod 4$$

genügt, so lassen sich die beiden Wurzeln der Kongruenz in der Form

$$x \equiv \pm a^{\frac{p+1}{4}} \mod p$$

darstellen. Dies folgt einfach daraus, daſs man die Kongruenz wegen $a^{\frac{p-1}{2}} \equiv 1$ mod p immer in der Form

$$x^2 \equiv a^{\frac{p+1}{2}} \mod p$$

schreiben kann, die wenn $\frac{p+1}{2}$ durch 2 teilbar ist, weiter in

$$\left(x - p^{\frac{p+1}{4}}\right)\left(x + a^{\frac{p+1}{4}}\right) \equiv 0 \mod p$$

übergeführt werden kann.

---

§ 75. Lösung der quadratischen Kongruenzen.

Die Lösung einer Kongruenz nach einem beliebigen Modul ist in § 68 auf die Lösung aller in diesem Modul enthaltenen Primzahlpotenzen zurückgeführt worden. Für die quadratischen Kongruenzen von der Form

$$x^2 \equiv a \mod m$$

haben wir aber erst den Fall behandelt, dafs der Modul m eine Primzahl ist. Um daher die Theorie dieser Kongruenz zu erledigen, müssen wir nun noch den Fall betrachten, dafs der Modul m eine Primzahlpotenz ist. Wir gehen hierbei von folgenden Betrachtungen aus.

Angenommen, man hätte die Kongruenz

$$x^2 \equiv a$$

nach zwei Moduln m_1 und m_2 gelöst, und nach ersterer das Lösungssystem $x \equiv \pm \alpha_1$, nach dem zweiten das Lösungssystem $x \equiv \pm \alpha_2$ gefunden, so dafs also

$$\alpha_1^2 \equiv a \mod m_1, \qquad \alpha_2^2 \equiv a \mod m_2$$

ist. Suchen wir nun eine Lösung nach dem Modul $m_1 m_2$, so mufs diese offenbar den Kongruenzen

$$x \equiv \alpha_1 \mod m_1, \qquad x \equiv \alpha_2 \mod m_2,$$

genügen. Beachten wir, dafs

$$x^2 = (x - \alpha_1)(x - \alpha_2) + (\alpha_1 + \alpha_2)x - \alpha_1 \alpha_2,$$

also

$$x^2 \equiv (\alpha_1 + \alpha_2)x - \alpha_1 \alpha_2 \mod m_1 m_2$$

ist, so kann man die Kongruenz $x^2 \equiv a \mod m_1 m_2$, die x auch befriedigen mufs, durch

$$(\alpha_1 + \alpha_2)x - \alpha_1 \alpha_2 \equiv a \mod m_1 m_2$$

ersetzen und somit das Problem auf die Lösung des folgenden Systems von linearen Kongruenzen zurückführen:

$$x \equiv \alpha_1 \mod m_1, \quad x \equiv \alpha_2 \mod m_2$$
$$(\alpha_1 + \alpha_2)x \equiv a + \alpha_1 \alpha_2 \mod m_1 m_2.$$

Schreiben wir dies in Form einer Äquivalenz

$$[m_1(x - \alpha_2),\ m_2(x - \alpha_1),$$
$$(\alpha_1 + \alpha_2)x - (a + \alpha_1 \alpha_2),\ m_1 m_2] = m_1 m_2,$$

so erkennen wir nach § 30 mit Anwendung einiger einfacher Reduktionen als hinreichende und notwendige Bedingung für die Lösbarkeit des Problems die folgende:

$$\left(m_1,\ m_2,\ \alpha_1 - \alpha_2,\ \alpha_1 + \alpha_2,\ \frac{\alpha_1^2 - a}{m_1},\ \frac{\alpha_2^2 - a}{m_2}\right)$$
$$= (m_1,\ m_2,\ \alpha_1 + \alpha_2).$$

Wenn nun $(a, m_1) = 1$ und m_1 als ungerade Zahl vorausgesetzt wird, so hat die linke Seite den Wert 1, da dann $(m_1,\ \alpha_1 - \alpha_2,\ \alpha_1 + \alpha_2) = (m_1,\ 2\alpha_1,\ \alpha_1 + \alpha_2) = 1$ wird; es mufs daher

$$(m_1,\ m_2,\ \alpha_1 + \alpha_2) = 1$$

sein. Benutzen wir nun die Äquivalenz (§ 20)

$$(a,\ b,\ c)\ (b\,c,\ c\,a,\ a\,b) = (b,\ c)\ (c,\ a)\ (a,\ b),$$

indem wir

$$a = (\alpha_1 - \alpha_2,\ m_1),\quad b = (\alpha_1 + \alpha_2,\ m_2),\quad c = (m_1,\ m_2)$$

setzen, so erhalten wir

$$(\alpha_1^2 - \alpha_2^2,\ m_1,\ m_2) = (\alpha_1 - \alpha_2,\ m_1,\ m_2)\ (\alpha_1 + \alpha_2,\ m_1,\ m_2).$$

Nun ist aber leicht einzusehen, dafs die linke Seite gleich (m_1, m_2) ist, da $\alpha_1^2 - a$, $\alpha_2^2 - a$ durch (m_1, m_2) teilbar ist und somit auch die Differenz $\alpha_1^2 - \alpha_2^2$. Es mufs somit, wenn (m_1, m_2) eine Primzahlpotenz ist, eins der beiden Modulsysteme

$$(\alpha_1 - \alpha_2,\ m_1,\ m_2),\quad (\alpha_1 + \alpha_2,\ m_1,\ m_2)$$

äquivalent (m_1, m_2), das andere äquivalent 1 sein.

Da nun zugleich mit α_2 auch $-\alpha_2$ eine Wurzel der Kongruenz $x^2 \equiv a \bmod m_2$ ist, so kann man durch einen etwa nötigen Vorzeichenwechsel stets erreichen, dafs das erstere Modulsystem äquivalent (m_1, m_2), das zweite äquivalent 1 wird. Daraus folgt dann die Lösbarkeit der Kongruenz

$$x^2 \equiv a \bmod m_1\, m_2,$$

und zwar ergiebt sich ein einziges Lösungssystem nach dem Modul $m_1\, m_2$.

Von den Folgerungen, die sich aus dieser Betrachtung ergeben, wollen wir nur diejenigen erwähnen, die für uns

nach unseren bisherigen Entwickelungen nur noch allein ein Interesse darbieten, und die sich auf die Lösung der Kongruenz $x^2 \equiv a$ nach der Potenz einer Primzahl als Modul beziehen. Sei p eine ungerade Primzahl, und nehmen wir $m_1 = m_2 = p$ an, so erhalten wir sofort, dafs jedem Lösungssystem der Kongruenz

$$x^2 \equiv a \mod p$$

ein solches der Kongruenz

$$x^2 \equiv a \mod p^2$$

entspricht.

Da nun aber die erstere zwei Lösungen besitzt, wenn $a^{\frac{p-1}{2}} \equiv 1 \mod p$ ist, so hat die letztere dann immer auch zwei Lösungssysteme.

Nehmen wir ferner $m_1 = p$, $m_2 = p^2$ an, so finden wir, dafs auch die Kongruenz

$$x^2 \equiv a \mod p^3$$

immer zwei Lösungen besitzt. So kann man weiter schliefsen und als Resultat folgenden Satz aussprechen:

Ist p eine ungerade Primzahl, zu der a teilerfremd ist, n eine beliebige ganze Zahl, so hat die Kongruenz

$$x^2 \equiv a \mod p^n$$

stets zwei und nur zwei Lösungssysteme nach dem Modul p^n, wenn a quadratischer Rest der Primzahl p ist.

Es bleibt jetzt noch übrig, Kongruenzen von der Form

$$x^2 \equiv a \mod 2^n$$

in Betracht zu ziehen, wo a zu 2 teilerfremd angenommen wird. Für $n = 1, 2, 3$ hat die Kongruenz ein Lösungssystem nach dem Modul 2,

$$x \equiv 1 \mod 2,$$

und zwar immer, wenn $n = 1$ ist; für $n = 2$ tritt die Bedingung $a \equiv 1 \mod 4$ und für $n = 3$ die Bedingung $a \equiv 1 \mod 8$ hinzu.

Man erkennt dies daraus, dafs das Quadrat jeder ungeraden Zahl, die sich in den Formen $2k + 1$, $4k \pm 1$

darstellen läfst, gleich $4k^2 + 4k + 1$, $16k^2 \pm 8k + 1$ ist, also der Einheit kongruent ist nach 4 und 8 als Modul. Wenn $n > 3$ angenommen wird, so läfst sich nun zeigen, dafs die Bedingung $a \equiv 1 \bmod 8$ genügt, um die Lösbarkeit der Kongruenz zu sichern, und dafs sie dann stets zwei Lösungssysteme nach dem Modul 2^{n-1} hat.

Dies beweisen wir durch vollständige Induktion, indem wir zeigen, dafs jeder Wurzel $x \equiv \alpha \bmod 2^{n-1}$ der Kongruenz $x^2 \equiv a \bmod 2^n$ eine Wurzel nach dem Modul 2^n der Kongruenz $x^2 \equiv a \bmod 2^{n+1}$ entspricht. Aus $x \equiv \alpha \bmod 2^{n-1}$ folgt $x^2 \equiv 2\alpha x - \alpha^2 \bmod 2^{2n-2}$, also auch mod 2^{n+1}, wenn $n \geqq 3$ ist. Die Kongruenz $x^2 \equiv a \bmod 2^n$ kann also ersetzt werden durch die folgende

$$2\alpha x - \alpha^2 \equiv a \bmod 2^{n+1}.$$

Da aber $\alpha^2 - a$ durch 2^n teilbar und somit $\alpha^2 + a$ gerade ist, so kann diese auch in

$$\alpha x \equiv \frac{a + \alpha^2}{2} \bmod 2^n$$

umgewandelt werden, die, weil α ungerade, stets lösbar ist. Ändert man das Vorzeichen von α, so erhält man die andere Wurzel der Kongruenz $x^2 \equiv a \bmod 2^n$ nach dem Modul 2^{n-1}, und es ist leicht zu sehen, dafs dieser Wurzel eine solche von $x^2 \equiv a \bmod 2^{n+1}$ entspricht, die durch Vorzeichenänderung der ersten hervorgeht.

Legen wir bei den Lösungssystemen immer denselben Modul zu Grunde wie bei den Kongruenzen, so folgt, dafs die Kongruenz

$$x^2 \equiv a \bmod 2^n$$

bei $n = 1$, $a \equiv 1 \bmod 2$ eine, bei $n = 2$, $a \equiv 1 \bmod 4$ zwei und bei $n \geqq 3$, $a \equiv 1 \bmod 3$ vier Lösungssysteme besitzt.

Jetzt sind wir vollständig ausgerüstet, die Bedingungen der Lösbarkeit und die Anzahl der Lösungssysteme der Kongruenz

$$x^2 \equiv a \bmod m \qquad ((a, m) = 1)$$

angeben zu können. Ist die Zerlegung von m in Primzahlen

$$m = 2^{e_0} p_1^{e_1} p_2^{e_2} \dots p_n^{e_n},$$

so ist die Kongruenz lösbar, wenn $a^{\frac{p_i - 1}{2}} \equiv 1 \mod p_i$, und hierzu tritt $a \equiv 1 \mod 2, \equiv 1 \mod 4, \equiv 3 \mod 8$, je nachdem $e_0 = 1, 2$ oder $\geqq 3$ ist. Aus § 68 und den obigen Untersuchungen folgt dann weiter, dafs die Anzahl der Lösungssysteme gleich 2^n, 2^{n+1}, 2^{n+2} ist, je nachdem $e_0 = 0, 1$ oder $e_0 = 2$ oder $e_0 \geqq 3$ ist.

Zum Schlufs wollen wir noch bemerken, dafs wir mit unseren Mitteln auch imstande sind, die gemischt quadratische Kongruenz

$$ax^2 + bx + c \equiv 0 \mod m$$

zu lösen, da diese der folgenden

$$4a^2x^2 + 4abx + 4ac \equiv 0 \mod 4am$$

äquivalent ist, die ihrerseits in der Form

$$(2ax + b)^2 \equiv b^2 - ac \mod 4_a m$$

geschrieben werden kann.

Beispiele: 1) Um die Kongruenz $x^2 \equiv 39 \mod 49$ zu lösen, bemerken wir, dafs aus ihr $x^2 \equiv 4 \mod 7$, $x \equiv \pm 2 \mod 7$ folgt. Man hat nun mit $m_1 = m_2 = 7$, $\alpha_1 = \alpha_2 = 2$, $a = 39$ die Äquivalenz

$$[7(x - 2),\ 4x - 43,\ 49] = 49$$

und erhält aus ihr durch Reduktion $x \equiv 23 \mod 49$. Die vorgelegte Kongruenz hat also die beiden Lösungsysteme $x \equiv 23, 26 \mod 49$.

2) Bei $x^2 \equiv 48 \mod 169$ ergiebt sich ebenso $x^2 \equiv 9 \mod 13$, $x \equiv \pm 3 \mod 13$. Es entsteht die Äquivalenz

$$[13(x - 3),\ 6x - 57,\ 169] = 169$$

und hieraus $x \equiv 94 \mod 169$. Die beiden Lösungssysteme sind also $x \equiv 75, 94 \mod 169$.

IX. Abschnitt.

Resultanten, Discriminanten und Elimination.

---

§ 76. Resultante zweier ganzer Funktionen.

Bei der Bestimmung des gröfsten gemeinschaftlichen Teilers zweier ganzen Funktionen (§ 62) haben wir ein Verfahren kennen gelernt, aus dem sich die Bedingungen für die Koeffizienten ermitteln lassen, damit die beiden Funktionen ohne gemeinschaftlichen Teiler sind. Es ergiebt sich nämlich durch das Euklidische Verfahren zuletzt eine einzige von den Koeffizienten beider Funktionen abhängige Gröfse, die nicht verschwinden darf. Doch ist es schwer, die allgemeine Form und die Eigenschaften dieser Gröfse hieraus abzuleiten, aber mit Hülfe der Determinantentheorie läfst sich dies in einfacher Weise durchführen.

Die beiden gegebenen ganzen Funktionen $A(x)$ und $B(x)$ seien dargestellt in der Form

$$A(x) = a_0 x^\alpha + a_1 x^{\alpha-1} + \ldots + a_\alpha$$
$$B(x) = b_0 x^\beta + b_1 x^{\beta-1} + \ldots + b_\beta.$$

Betrachten wir nun die Funktionen

$$A(x),\ A(x)x, \ldots A(x)x^{\beta-1}$$
$$B(x),\ B(x)x, \ldots B(x)x^{\alpha-1},$$

so lassen sich diese als ein System von $(\alpha + \beta)$ in den $(\alpha + \beta)$ Gröfsen

$$1,\ x,\ x^2, \ldots x^{\alpha+\beta-1}$$

homogenen linearen Funktionen auffassen. Wenn nun $A(x)$ und $B(x)$ teilerfremd sind, so kann man die Zahl 1 und

überhaupt jede Potenz von x homogen und linear durch A(x) und B(x) ausdrücken mit Koeffizienten, die ganze Funktionen von x sind. Die hinreichende und notwendige Bedingung dafür, dafs sich die genannten $(\alpha + \beta)$ Gröfsen linear durch die betrachteten $(\alpha + \beta)$ Funktionen ausdrücken besteht darin, dafs die Determinante des Funktionensystems nicht verschwindet. Diese Determinante ist aber die folgende

$$\begin{vmatrix} a_0 & a_1 \ldots a_\alpha & 0 \ldots 0 \\ 0 & a_0 \ldots a_{\alpha-1} & a_\alpha \ldots 0 \\ \vdots & & \\ 0 & 0 \ldots a_0 & a_1 \ldots a_\alpha \\ b_0 & b_1 \ldots b_\beta & 0 \ldots 0 \\ 0 & b_0 \ldots b_{\beta-1} & b_\beta \ldots 0 \\ \vdots & & \\ 0 & 0 \ldots b_0 & b_1 \ldots b_\beta \end{vmatrix}.$$

Sie wird die **Resultante** der beiden Funktionen A(x) und B(x) genannt und mit $\mathbf{R}(A(x), B(x))$ bezeichnet.

Die hinreichende und notwendige Bedingung dafür, dafs zwei ganze Funktionen A(x) und B(x) teilerfremd sind, besteht darin, dafs ihre Resultante $\mathbf{R}(A, B) \neq 0$ ist.

Danach ergiebt sich als Resultante der beiden linearen Funktionen

$$a_0 x + a_1, \qquad b_0 x + b_1$$

einfach die Determinante

$$\begin{vmatrix} a_0 & a_1 \\ b_0 & b_1 \end{vmatrix},$$

so dafs der Resultantenbegriff als eine Erweiterung des Begriffes der Determinante angesehen werden kann.

Eine sehr einfache Rechnung lehrt ferner, dafs

$$\mathbf{R}[x, A(x)] = A(0) = a_\alpha$$
$$\mathbf{R}[x - a, A(x)] = A(a)$$

ist in Übereinstimmung mit den früher abgeleiteten Sätzen, dafs eine Funktion nur durch ihre Variable teilbar ist, wenn ihr konstantes Glied fehlt, und allgemein durch einen linearen Faktor teilbar ist, wenn sie zugleich mit diesem verschwindet (§ 59).

§ 77. Eigenschaften der Resultanten hinsichtlich der Funktionen.

I. Bei der Vertauschung der Funktionen gilt das Gesetz

$$\mathbf{R}(A, B) = (-1)^{\alpha\beta}\,\mathbf{R}(B, A).$$

Um nämlich von $\mathbf{R}(A, B)$ zu $\mathbf{R}(B, A)$ überzugehen, hat man in der Determinante von $\mathbf{R}(A, B)$ nur eine Vertauschung der Zeilen vorzunehmen, nämlich von der Anordnung

$$1,\ 2, \ldots \beta,\ \beta + 1,\ \beta + 2, \ldots \alpha + \beta$$

zu der Anordnung

$$\beta + 1,\ \beta + 2, \ldots \beta + \alpha,\ 1,\ 2, \ldots \beta$$

überzugehen. Wie man leicht erkennt, kann dies geschehen durch successive Anwendung der Permutationen

$$(\beta + 1, \beta, \ldots 2, 1),\ (\beta + 2, \beta, \ldots 2, 1,), \ldots (\beta + \alpha, \beta, \ldots 2, 1)$$

Jeder von diesen α Cyklen von der Ordnung $\beta + 1$ läfst sich in β Transpositionen zerlegen (§ 39), daher ist die Permutation durch $\alpha\beta$ Transpositionen ersetzbar und mufs die Determinante um das Vorzeichen $(-1)^{\alpha\beta}$ ändern.

II. Dafs die Resultante $\mathbf{R}(A, B)$ eine homogene Funktion der Koeffizienten a und b, in ersteren vom Grade β, in letzteren vom Grade α ist, ergiebt sich unmittelbar aus der Form der Determinantendarstellung. Wir drücken diese Eigenschaften aus durch die Gleichungen:

$$\mathbf{R}(tA, B) = t^{\beta}\,\mathbf{R}(A, B)$$
$$\mathbf{R}(A, tB) = t^{\alpha}\,\mathbf{R}(A, B),$$

von denen sich mit Hülfe von I die eine aus der andern ableiten läfst, und aus denen noch folgt

$$\mathbf{R}(t\,A, t\,B) = t^{\alpha+\beta}\,\mathbf{R}(A, B).$$

III. Ist der Grad α von $A(x)$ gröfser als der Grad β von $B(x)$, $C(x)$ von einem Grade $\gamma \leqq \alpha - \beta$, so ist

$$\mathbf{R}(A + BC, B) = \mathbf{R}(A, B).$$

Wir bezeichnen die Koeffizienten von $C(x)$ mit c_0, $c_1, \ldots c_\gamma$ und multiplizieren, wenn i eine der Zahlen $1, 2, \ldots \beta$ bedeutet, die Elemente der $(\alpha - \gamma + i)$ten, $(\alpha - \gamma + i - 1)$ten $\ldots (\alpha + i)$ten Zeile in der Determinante für $\mathbf{R}(A, B)$ mit

$c_0, c_1, \dots c_\gamma$ und summieren sodann zur $i$ten Zeile; dann werden die Elemente dieser $i$ten Zeile $a_{\alpha-\beta-\gamma}, a_{\alpha-\beta-\gamma+1}, \dots a_\alpha$ vermehrt um die Gröfsen $b_0 c_0$, $b_0 c_1 + b_1 c_0, \dots b_\beta c_\gamma$, die in der Entwickelung von

$$B(x)\,C(x) = b_0 c_0 x^{\beta+\gamma} + (b_0 c_1 + b_1 c_0) x^{\beta+\gamma-1} + \dots + b_\beta c_\gamma$$

auftreten, es entsteht also $\mathbf{R}(A + BC, B)$.

Aus diesem Satze läfst sich eine wichtige Folgerung ziehen, die zu einer Berechnung der Resultanten führt. Sind $A(x)$ und $B(x)$ wieder zwei ganze Funktionen, und ist $\alpha > \beta$, so kann man (§ 56) zwei ganze Funktionen $G(x)$ und $C(x)$ so bestimmen, dafs

$$A(x) = G(x)\; B(x) + C(x)$$

und der Grad γ von $C(x)$ kleiner als der Grad β von $B(x)$ ist. Da der Grad von $G(x)\,B(x)$ gleich α ist, so giebt die Anwendung des soeben bewiesenen Satzes

$$\begin{aligned}\mathbf{R}(A, B) &= \mathbf{R}(A - GB, B) = \mathbf{R}(C, B) \\ &= (-1)^{\beta\gamma}\,\mathbf{R}(B, C).\end{aligned}$$

Mit $\mathbf{R}(B, C)$ kann man wieder verfahren wie mit $\mathbf{R}(A, B)$ und so die bei der Aufsuchung des gröfsten gemeinschaftlichen Teilers (§ 62) auftretenden Funktionen benutzen, um die Resultante zu berechnen.

IV. Multiplikationstheorem des Resultanten:

$$\mathbf{R}(AB, C) = \mathbf{R}(A, C)\,\mathbf{R}(B, C).$$

Sind $a_0, a_1, \dots a_\alpha$ die Koeffizienten von $A(x)$, $b_0, b_1, \dots b_\beta$ die von $B(x)$ und $c_0, c_1, \dots c_\gamma$ die von $C(x)$, so ist

$$\mathbf{R}(A, C) = \begin{vmatrix} a_0 & a_1 & \dots & a_\alpha & & \\ & \ddots & & & & \\ & & a_0 & a_1 & \dots & a_\alpha \\ c_0 & c_1 & \dots & c_\gamma & & \\ & \ddots & & & & \\ & & c_0 & c_1 & \dots & c_\gamma \end{vmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} (\gamma \text{ Zeilen}) \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} (\alpha \text{ Zeilen}) \end{matrix}$$

$$\mathbf{R}(B, C) = \begin{vmatrix} b_0 & b_1 & \dots & b_\beta & & \\ & \ddots & & & & \\ & & b_0 & b_1 & \dots & b_\beta \\ c_0 & c_1 & \dots & c_\gamma & & \\ & \ddots & & & & \\ & & c_0 & c_1 & \dots & c_\gamma \end{vmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} (\gamma \text{ Zeilen}) \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} (\beta \text{ Zeilen}). \end{matrix}$$

Beide Determinanten sind von verschiedenem Grade, lassen sich aber als Determinanten vom Grade $(\alpha + \beta + \gamma)$

darstellen, wenn wir die Diagonale der ersten um β Elemente 1 fortsetzen, in der zweiten links oben die Reihe $b_0, b_1, .. b_\beta$ α mal ansetzen (s. § 50). So erhalten wir

$$\mathbf{R}(A, C) = \left| \begin{matrix} a_0 & a_1 & \ldots & a_\alpha & & & 0 \\ & \ddots & & & & & \\ & & a_0 & a_1 & \ldots & a_\alpha \ldots & 0 \\ c_0 & c_1 & \ldots & c_\gamma & \ldots & \ldots & 0 \\ & \ddots & & & & & \\ & & c_0 & c_1 & \ldots & c_\gamma & \\ 0 & \ldots & \ldots & \ldots & 1 & \ldots & 0 \\ \vdots & & & & & \ddots & \\ 0 & \ldots & \ldots & \ldots & 0 & \ldots & 1 \end{matrix} \right| \begin{matrix} \} (\gamma \text{ Zeilen}) \\ \\ \} (\alpha \text{ Zeilen}) \\ \\ \} (\beta \text{ Zeilen}) \end{matrix}$$

$$b_0^\alpha \mathbf{R}(B, C) = \left| \begin{matrix} b_0 & b_1 & \ldots & b_\beta & & & \\ & \ddots & & & & & \\ & b_0 & b_1 & \ldots \ldots & b_\beta & & \\ & & \ddots & & & & \\ & & & b_0 & b_1 & \ldots \ldots & b_\beta \\ & c_0 & c_1 & \ldots \ldots & c_\gamma & & \\ & & \ddots & & & & \\ & & & c_0 & c_1 & \ldots \ldots & c_\gamma \end{matrix} \right| \begin{matrix} \} ((\alpha + \gamma) \text{ Zeilen}) \\ \\ \} (\beta \text{ Zeilen}). \end{matrix}$$

Bilden wir nun das Produkt mit Anwendung des Multiplikationstheorems der Determinanten, so erhalten wir, wenn wir mit $d_0, d_1, \ldots d_{\alpha+\beta}$ die Koeffizienten von $A(x)\,B(x)$, mit $e_0, e_1, \ldots e_{\beta+\gamma}$ die von $B(x)\,C(x)$ bezeichnen,

$$b_0^\alpha \mathbf{R}(A, C)\,\mathbf{R}(B, C) = \left| \begin{matrix} d_0 & d_1 & \ldots & d_{\alpha+\beta} & & \\ & \ddots & & & & \\ & d_0 & d_1 & \ldots \ldots \ldots & d_{\alpha+\beta} & \\ e_0 & e_1 & \ldots & e_{\beta+\gamma} & & \\ & \ddots & & & & \\ & e_0 & e_1 & \ldots \ldots \ldots & e_{\beta+\gamma} & \\ & & c_0 & \ldots \ldots & c_\gamma & \\ & & & \ddots & & \\ & & & c_0 & \ldots & c_\gamma \end{matrix} \right| \begin{matrix} \} (\gamma \text{ Zeilen}) \\ \\ \} (\alpha \text{ Zeilen}). \\ \\ \} (\beta \text{ Zeilen}) \end{matrix}$$

Multiplizieren wir die Elemente der $(\gamma + \alpha + 1)$ten, $(\gamma + \alpha + 2)$ten, ... $(\gamma + \alpha + \beta)$ten Zeile resp. mit b_1, $b_2, \ldots b_\beta$, und subtrahieren von der $(\gamma + \alpha)$ten Zeile, so treten an Stelle der Elemente $e_0, e_1, \ldots e_{\beta+\gamma}$ die Produkte $b_0 c_0, b_0 c_1, \ldots b_0 c_\gamma$ und weiterhin lauter Nullen. Wir können aus der $(\alpha + \gamma)$ten Zeile die Gröſse b_0 herausheben. Ist das geschehen, so multiplizieren wir die Elemente der $(\gamma + \alpha)$ten bis $(\gamma + \alpha + \beta - 1)$ten Zeile resp. mit $b_1, b_2, \ldots b_\beta$ und verfahren wie vorhin. Nach α solchen Operationen

erhalten wir nach Forthebung des auf beiden Seiten stehenden Faktors b_0^α

$$\mathbf{R}(A, C)\,\mathbf{R}(B, C) = \begin{vmatrix} d_0 & d_1 & \dots & d_{\alpha+\beta} & & \\ & \ddots & & & & \\ & & d_0 & \dots & d_{\alpha+\beta-1} & d_{\alpha+\beta} \\ c_0 & c_1 & \dots & c_\gamma & & \\ & \ddots & & & & \\ & & c_0 & c_1 & \dots & c_\gamma \end{vmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} (\gamma \text{ Zeilen}) \\ \left.\vphantom{\begin{matrix} a \\ a \\ a \end{matrix}}\right\} (\alpha+\beta) \text{ Zeilen}), \end{matrix}$$

und die rechts stehende Determinante ist offenbar gleich $\mathbf{R}(AB, C)$.

Das Multiplikationstheorem läfst sich leicht auf mehrere Faktoren erweitern. Um seine formale Gültigkeit auch für den Fall aufrecht zu erhalten, dafs an die Stelle der einen Funktion eine Konstante tritt, wo eine eigentliche Resultante nicht mehr existiert, hat man

$$\mathbf{R}(a, A(x) = a^\alpha$$

anzunehmen, weil dann einerseits

$$\mathbf{R}(a\,B(x), A(x)) = \mathbf{R}(a, A(x))\,\mathbf{R}(B, A)$$

sein soll, während andererseits nach (II)

$$\mathbf{R}(a\,B, A) = a^\alpha\,\mathbf{R}(B, A)$$

ist.

Zerfällt eine der beiden Funktionen $A(x)$ und $B(x)$ oder beide in Linearfaktoren

$$\begin{aligned} A(x) &= a_0\,(x - a_1)\,(x - a_2)\dots(x - a_\alpha) \\ B(x) &= b_0\,(x - b_1)\,(x - b_2)\dots(x - b_\beta), \end{aligned}$$

so ergiebt sich mit wiederholter Anwendung des Multiplikationstheorems

$$\begin{aligned} \mathbf{R}(A, B) &= a_0^\beta \prod_i B(a_i) = (-1)^{\alpha\beta}\, b_0^\alpha \prod_k A(b_k) \\ &= a_0^\beta\, b_0^\alpha \prod_{i,k} (a_i - b_k). \end{aligned}$$

$$(i = 1, 2, \dots \alpha;\ k = 1, 2, \dots \beta)$$

Diese Ausdrücke werden häufig als Ausgangspunkt für die Theorie der Resultanten benutzt, wenn man die Theorie der symmetrischen Funktionen voraussetzt, aus denen sich dann eine Methode zu ihrer Berechnung ergiebt (§ 100).

Aus den Sätzen (II, III, IV) leitet man leicht folgende Gleichung ab, wenn $A(x)$ und $B(x)$ denselben Grad n haben:

$$\mathbf{R}(aA + bB, cA + dB) = (ad - bc)^n\,\mathbf{R}(A, B).$$

Es ergiebt sich successive

$$\begin{aligned}
&\mathbf{R}(aA + bB, cA + dB) = d^{-n}\mathbf{R}(adA + bdB, cA + dB)\\
&= d^{-n}\mathbf{R}((ad - bc)A, cA + dB)\\
&= d^{-n}(ad - bc)^n\mathbf{R}(A, cA + dB)\\
&= d^{-n}(ad - bc)^n\mathbf{R}(A, dB) = d^{-n}(ad - bc)\, d^{-n}\mathbf{R}(A, B)\\
&= (ad - bc)^n\mathbf{R}(A, B).
\end{aligned}$$

Hierbei ist vorausgesetzt, dafs $d \neq 0$ ist; man kann die Ableitung aber leicht von dieser Annahme befreien, wenn nur $ad - bc \neq 0$ ist.

---

§ 78. Verhalten der Resultante bei linearer Transformation der Variabeln.

Unterwirft man die Variable x der Funktionen $A(x)$ und $B(x)$ einer linearen Transformation und bildet von den so entstehenden Funktionen die Resultante, so steht diese in einfacher Beziehung zur Resultante $\mathbf{R}(A(x), B(x))$ der ursprünglichen Funktionen. Wir gehen schrittweise vor und untersuchen erst einige spezielle Fälle, um dann das allgemeine Resultat abzuleiten.

Multipliziert man die Variable x mit einer konstanten Gröfse t, so ergiebt sich

$$\text{(I)} \qquad \mathbf{R}(A(tx), B(tx)) = t^{\alpha\beta}\,\mathbf{R}(A(x), B(x)).$$

Denn die Determinante für $\mathbf{R}(A(tx), B(tx))$

$$\begin{vmatrix}
a_0 t^\alpha & a_1 t^{\alpha-1} \dots a_\alpha & & \\
\ddots & & \ddots & \\
& a_0 t^\alpha & a_1 t^{\alpha-1} \dots & a_\alpha \\
b_0 t^\beta & b_1 t^{\beta-1} \dots b_\beta & & \\
\ddots & & \ddots & \\
& b_0 t^\beta & b_1 t^{\beta-1} \dots & b_\beta
\end{vmatrix}$$

kann man aus der Determinante für $\mathbf{R}(A(x), B(x)$ dadurch ableiten, dafs man in dieser die Elemente der Zeilen der Reihe nach mit $t^{\alpha+\beta-1}, t^{\alpha+\beta-2}, \dots t^2, t, 1$ multipliziert und darauf die Elemente der Spalten durch $t^{\beta-1}, \dots t, 1, t^{\alpha-1}, \dots t, 1$ dividiert. Dadurch wird die Determinante $\mathbf{R}(A(x), B(x))$ aber um eine Potenz von t mit dem Exponenten

$$\sum_{i=0}^{i=\alpha+\beta-1} i - \sum_{i=0}^{i=\alpha-1} i - \sum_{i=0}^{i=\beta-1} i$$

$$= \frac{(\alpha+\beta)(\alpha+\beta-1)}{2} - \frac{\alpha(\alpha-1)}{2} - \frac{\beta(\beta-1)}{2} = \alpha\beta$$

geändert.

Vermehrt man die Variable x um die Einheit, so zeigt sich, dafs

$$\text{(II)} \quad R(A(x+1), B(x+1)) = R(A(x), B(x))$$

ist. Doch ist der Beweis nicht ganz einfach.

Wir bemerken dazu zunächst, dafs $R(A(x+1), B(x+1))$ aus $R(A(x), B(x))$ hervorgeht, wenn man an Stelle von $a_0, a_1, \ldots a_\alpha, b_0, b_1, \ldots b_\beta$ die Gröfsen $a_0', a_1', \ldots a_\alpha', b_0', b_1', \ldots b_\beta'$ setzt, die in den Entwickelungen

$$A(x+1) = a_0' x^\alpha + a_1' x^{\alpha-1} + \ldots + a_\alpha'$$
$$B(x+1) = b_0' x^\beta + b_1' x^{\beta-1} + \ldots + b_\beta'$$

auftreten und allgemein durch die folgenden Gleichungen bestimmt sind:

$$a_i' = \sum_h (\alpha - h)_{\alpha - i}\, a_h, \quad b_i' = \sum_h (\beta - h)_{\beta - i}\, b_h. \qquad (h = 0, 1, \ldots i)$$

Wir komponieren nun die Matrix

$$\begin{pmatrix} (\beta-1)_{\beta-1} & (\beta-1)_{\beta-2} \ldots (\beta-1)_0 & & & \\ & (\beta-2)_{\beta-2} \ldots (\beta-2)_0 & & & \\ & \ddots \quad \vdots & & & \\ & 1 & & & \\ & & (\alpha-1)_{\alpha-1} & (\alpha-1)_{\alpha-2} \ldots (\alpha-1)_0 \\ & & & (\alpha-2)_{\alpha-2} \ldots (\alpha-2)_0 \\ & & & \ddots \quad \vdots \\ & & & 1 \end{pmatrix},$$

deren Determinante den Wert 1 hat, mit der Matrix

$$\begin{pmatrix} a_0' & a_1' \ldots a_\alpha' & & \\ & a_0' \ldots a'_{\alpha-1} & a_\alpha' & \\ & \ddots & & \\ & & a_0' \ldots\ldots\ldots & a_\alpha' \\ b_0' & b_1' \ldots b_\beta' & & \\ & b_0' \ldots b'_{\beta-1} & b_\beta' & \\ & \ddots & & \\ & & b_0' \ldots\ldots\ldots & b_\beta' \end{pmatrix}$$

und nennen die entstehende Matrix

$$(c_{ik}). \qquad (i, k = 0, 1, \ldots \alpha + \beta - 1)$$

Die β ersten Zeilen entstehen also aus der Komposition der Systeme

$$\begin{pmatrix} (\beta - i - 1)_{\beta - k - 1} \\ (i, k = 0, 1, \ldots \beta - 1) \end{pmatrix} \text{ und } \begin{pmatrix} a'_{k-i} \, \begin{matrix} i = 0, 1, \ldots \beta - 1; \\ k = 0, 1, \ldots \alpha + \beta - 1) \end{matrix} \end{pmatrix},$$

die α letzten Zeilen aus der Komposition von

$$\begin{pmatrix} (\alpha - i - 1)_{\alpha - k - 1} \\ (i, k = 0, 1, \ldots \alpha - 1) \end{pmatrix} \text{ mit } \begin{pmatrix} b'_{k-i} \, \begin{matrix} (i = 0, 1, \ldots \alpha - 1; \\ k = 0, 1, \ldots \alpha + \beta - 1) \end{matrix} \end{pmatrix}.$$

Hieraus ermittelt man leicht die Werte von c_{ik}; es ist nämlich für $i = 0, 1, \ldots \beta - 1$

$$c_{ik} = \sum_h (\beta - i - 1)_{\beta - h - 1} \, a'_{k-h},$$

dagegen für $i = \beta, \beta + 1, \ldots \beta + \alpha - 1$

$$c_{ik} = \sum_h (\alpha - i - 1)_{\alpha - h - 1} \, b'_{k-h}.$$

Setzt man für die Gröfsen a_i' und b_i' ihre oben angegebenen Werte ein und beachtet die für Binomialkoeffizienten geltende Formel

$$\sum_h i_h \, k_{r-h} = (i + k)_r, \qquad (h = 0, 1, \ldots r)$$

so ergiebt sich für $i = 0, 1, \ldots \beta - 1$

$$\begin{aligned} c_{ik} &= \sum_{h,g} (\beta - i - 1)_{\beta - h - 1} \, (\alpha - g)_{\alpha - k + h} \, a_g \\ &= \sum_g (\alpha + \beta - i - g - 1)_{\alpha + \beta - k - 1} \, a_g, \end{aligned}$$

für $\beta + i = \beta, \beta + 1, \ldots \beta + \alpha - 1$

$$\begin{aligned} c_{\beta + i\, k} &= \sum_h (\alpha - i - 1)_{\alpha - h - 1} \, (\beta - g)_{\beta - k + h} \, b_g \\ &= \sum_g (\alpha + \beta - i - g - 1)_{\alpha + \beta - k - 1} \, b_g. \end{aligned}$$

Wenn wir nun noch die Matrix $(c_{ik}, \; (i, k = 0, 1, \ldots \alpha + \beta - 1))$ mit der Matrix

$$\begin{pmatrix} (-1)^{i-k} \, (\alpha + \beta - i - 1)_{\alpha + \beta - k - 1} \\ (i, k = 0, 1, \ldots \alpha + \beta - 1) \end{pmatrix},$$

deren Determinante den Wert 1 hat, zusammensetzen, so erhalten wir die Matrix, deren Determinante $R(A(x), B(x))$ ist. Bezeichnen wir vorläufig das entstehende System mit $(d_{ik}, \ (i, k = 0, 1, \ldots \alpha + \beta - 1))$, so ist

$$d_{ik} = \sum_h (-1)^{h-k} c_{ih} \, (\alpha + \beta - h - 1)_{\alpha+\beta-k-1},$$

und es ergiebt sich mit Anwendung der Formel (§ 5, (10))

$$\sum_h (-1)^h i_h h_k = (-1)^k \delta_{ik} \quad (h = k, k+1, \ldots i-1, i)$$

für $i = 0, 1, \ldots \beta - 1$

$$\begin{aligned} d_{ik} &= \sum_{h,g} (-1)^{h-k} (\alpha + \beta - h - 1)_{\alpha+\beta-k-1} \\ &\qquad \times (\alpha + \beta - i - g - 1)_{\alpha+\beta-h-1} \, a_g \\ &= \sum_g a_g \sum_h (-1)^{h-k} (\alpha + \beta - h - 1)_{\alpha+\beta-k-1} \\ &\qquad \times (\alpha + \beta - i - g - 1)_{\alpha+\beta-h-1} \\ &= \sum_g \delta_{k, i+g} \, a_g = a_{k-i}, \end{aligned}$$

für $\beta + i = \beta, \beta + 1, \ldots \beta + \alpha - 1$ ebenso

$$\begin{aligned} d_{\beta+i\,k} &= \sum_{h,g} (-1)^{h-k} (\alpha + \beta - h - 1)_{\alpha+\beta-k-1} \\ &\qquad \times (\alpha + \beta - i - g - 1)_{\alpha+\beta-h-1} \, b_g \\ &= \sum_g b_g \sum_h (-1)^{h-k} (\alpha + \beta - h - 1)_{\alpha+\beta-k-1} \\ &\qquad \times (\alpha + \beta - i - g - 1)_{\alpha+\beta-h-1} \\ &= \sum_g \delta_{k, i+g} \, b_g = b_{k-i}. \end{aligned}$$

Wendet man nun noch das Multiplikationstheorem der Determinanten an, so erhält man das gewünschte Resultat.

Die hier angewandte Methode läfst sich mit einer geringfügigen Modifikation dazu benutzen, um die allgemeine Gleichung

$$\text{(III)} \quad \mathbf{R}(A(x + t), \ B(x + t)) = \mathbf{R}(A(x), \ B(x))$$

zu beweisen. Wir empfehlen dem Leser, dies auszuführen, schlagen aber selbst einen andern Weg ein, indem wir die Formel (I) benutzen; ihr zufolge ist nämlich

$$\mathbf{R}(A(tx), \ B(tx)) = t^{\alpha\beta} \, \mathbf{R}(A(x), \ B(x))$$

$$\mathbf{R}(A(tx + t), \ B(tx + t)) = t^{\alpha\beta} \, \mathbf{R}(A(x + t), \ B(x + t)),$$

während nach (II)

$$\mathbf{R}(A(t(x+1)),\ B(t(x+1))) = \mathbf{R}(A(tx),\ B(tx))$$

ist. Aus den hingeschriebenen Gleichungen folgt dann sofort die zu beweisende Relation.

Wir sind jetzt imstande, den Einfluſs zu bestimmen, den eine ganze lineare Transformation der Variabeln auf die Resultante hervorruft. Da nämlich

$$\mathbf{R}(A(ax+b), B(ax+b)) = \mathbf{R}\left[A\left(a\left(x+\frac{b}{a}\right)\right), A\left(a\left(x+\frac{b}{a}\right)\right)\right]$$
$$= \mathbf{R}(A(ax), B(ax))$$

$$\mathbf{R}(A(ax), B(ax)) = a^{\alpha\beta}\,\mathbf{R}(A(x), B(x))$$

ist, so folgt

$$\text{(IV)}\quad \mathbf{R}(A(ax+b), B(ax+b)) = a^{\alpha\beta}\,\mathbf{R}(A(x), B(x)).$$

Aber auch auf gebrochene lineare Transformationen läſst sich die Untersuchung ausdehnen. Zunächst betrachten wir wieder einen speziellen Fall und bemerken, daſs $x^\alpha A\left(\frac{1}{x}\right)$, $x^\beta B\left(\frac{1}{x}\right)$ ganze Funktionen sind, die die Form haben:

$$x^\alpha A\left(\frac{1}{x}\right) = a_\alpha x^\alpha + a_{\alpha-1} x^{\alpha-1} + \ldots + a_1 x + a_0$$

$$x^\beta B\left(\frac{1}{x}\right) = b_\beta x^\beta + b_{\beta-1} x^{\beta-1} + \ldots + b_1 x + b_0.$$

Ihre Resultante wird dargestellt durch die Determinante

$$\begin{vmatrix} a_\alpha & a_{\alpha-1} & \ldots & a_0 & & \\ & a_\alpha & \ldots & a_1 & a_0 & \\ & & \ddots & & & \\ & & a_\alpha & \ldots & \ldots & a_0 \\ b_\beta & b_{\beta-1} & \ldots & b_0 & & \\ & b_\beta & \ldots & b_1 & b_0 & \\ & & \ddots & & & \\ & & b_\beta & \ldots & \ldots & b_0 \end{vmatrix}.$$

Diese kann man aus der Determinante von $\mathbf{R}(A(x), B(x))$ auf folgende Weise durch Vertauschung der Zeilen und

Spalten erhalten. Man unterwerfe zunächst die Zeilen der Permutation

$$\begin{pmatrix} 1 & 2 & \dots \beta & \beta+1 & \beta+2 & \dots \beta+\alpha \\ \beta & \beta-1 & \dots 1 & \beta+\alpha & \beta+\alpha-1 & \dots \beta+1 \end{pmatrix},$$

was durch die Cyklen $(1, 2, \dots \beta)$, $(2, 3, \dots \beta)$, $\dots (\beta - 1, \beta)$; $(\beta + 1, \beta + 2, \dots \beta + \alpha)$, $(\beta + 2, \dots \beta + \alpha)$, $\dots (\beta + \alpha - 1, \beta + \alpha)$ geschehen kann und daher die Vorzeichenänderung $(-1)^{\frac{\alpha(\alpha-1)}{2}+\frac{\beta(\beta-1)}{2}}$ nach sich zieht. Hierauf sind nun noch die Spalten durch die Permutation

$$\begin{pmatrix} 1 & 2 & \dots \alpha+\beta \\ \alpha+\beta & a+\beta-1 & \dots 1 \end{pmatrix}$$

zu vertauschen, was durch die Cyklen $(1, 2, \dots \alpha + \beta)$, $(2, 3, \dots \alpha + \beta)$, $\dots (\alpha + \beta - 1, \alpha + \beta)$ erreicht werden kann und die Vorzeichenänderung $(-1)^{\frac{(\alpha+\beta)(\alpha+\beta-1)}{2}}$ zur Folge hat. Das Vorzeichen ändert sich also um

$$(-1)^{\frac{\alpha(\alpha-1)}{2}+\frac{\beta(\beta-1)}{2}+\frac{(\alpha+\beta)(\alpha+\beta-1)}{2}} = (-1)^{\alpha\beta},$$

und wir finden also, dafs

$$\text{(V)} \quad \mathbf{R}\left(x^\alpha t\left(\frac{1}{x}\right),\ x^\beta B\left(\frac{1}{x}\right)\right) = (-1)^{\alpha\beta}\,\mathbf{R}\,(A(x),\ B(x))$$

ist.

Alle vorhergehenden Formeln umfafst die folgende

$$\text{(VI)} \quad \mathbf{R}\left((cx+d)^\alpha A\left(\frac{ax+b}{cx+d}\right),\ (cx+d)^\beta B\left(\frac{ax+b}{cx+d}\right)\right) = (ad-bc)^{\alpha\beta}\,\mathbf{R}\,(A(x),\ B(x)),$$

die allgemein lehrt, wie sich die Resultante bei Anwendung einer allgemeinen gebrochenen linearen Substitution ändert, deren Determinante $ad - bc \neq 0$ sein mufs. Da dann nicht alle vier Elemente a, b, c, d verschwinden können, so mag etwa $c \neq 0$ angenommen werden. Dann läfst sich die lineare Substitution $\frac{ax+b}{cx+d}$ aus den folgenden zusammensetzen (§ 14): $x + \frac{a}{c}$, $-\frac{ad-bc}{c}x$, $\frac{1}{x}$, cx, $x + \frac{d}{c}$.

Die erste ändert die Resultante überhaupt nicht nach (III), die zweite nach (I) um den Faktor $\left(-\frac{a\,d-b\,c}{c}\right)^{\alpha\beta}$, die dritte nach (V) um das Vorzeichen $(-1)^{\alpha\beta}$, während die vierte den Faktor $c^{\alpha\beta}$ einführt, und die letzte keine Aenderung bewirkt. Die Gesamtänderung ist also

$$\left(-\frac{a\,d-b\,c}{c}\right)^{\alpha\beta}(-1)^{\alpha\beta}c^{\alpha\beta}=(a\,d-b\,c)^{\alpha\beta},$$

wie die Formel (VI) behauptet.

---

§ 79. Bau der Resultante hinsichtlich der Koeffizienten ihrer Funktionen.

Wir haben bereits in § 77 II bemerkt, dafs die Resultante $R(A(x), B(x)$ eine homogene Funktion der Koeffizienten a und b ist, in den ersteren vom Grade β, in den letzteren vom Grade α. Daher mufs sie sich in der Form darstellen lassen

$$R(A, B)=\Sigma\, c_{i_0 i_1 \dots i_\alpha k_0 k_1 \dots k_\beta} a_0^{i_0} a_1^{i_1} \dots a_\alpha^{i_\alpha} b_0^{k_0} b_1^{k_1} \dots b_\beta^{k_\beta},$$

wobei die Summation über die Indizes $i_0, i_1, \dots i_\alpha, k_0, k_1, \dots k_\beta$ zu erstrecken ist, und die Koeffizienten durch c bezeichnet worden sind, und es mufs sein

$$(1)\qquad \begin{aligned} i_0+i_1+\dots+i_\alpha&=\beta\\ k_0+k_1+\dots+k_\beta&=\alpha. \end{aligned}$$

Beachten wir nun noch die Gleichung $R(A(t\,x), B(t\,x)) = t^{\alpha\beta}R(A(x)\,B(x))$, und bedenken wir, dafs die Koeffizienten in $A(t\,x)$ und $B(t\,x)$ $a_0 t^\alpha$, $a_1 t^{\alpha-1}, \dots a_\alpha$, $b_0 t^\beta$, $b_1 t^{\beta-1}, \dots b_\beta$ sind, so ergiebt sich aus der Gleichung

$$\Sigma\, c_{i_0 i_1 \dots i_\alpha k_0 k_1 \dots k_\beta} a_0^{i_0} t^{\alpha i_0} a_1^{i_1} t^{(\alpha-1) i_1} \dots a_\alpha^{i_\alpha} b_0^{k_0} t^{\beta k_0} b_1^{k_1} t^{(\beta-1) k_1} \dots b_\beta^{k_\beta}$$
$$= t^{\alpha\beta}\, \Sigma\, c_{i_0 i_1 \dots i_\alpha k_0 k_1 \dots k_\beta} a_0^{i_0} a_1^{i_1} \dots a_\alpha^{i_\alpha} b_0^{k_0} b_1^{k_1} \dots b_\beta^{k_\beta},$$

dafs

$$(2')\ \alpha i_0+(\alpha-1)i_1+\dots+i_\alpha+\beta k_0+(\beta-1)k_1+\dots+i_\beta=\alpha\beta$$

ist. Mit Berücksichtigung der vorher erhaltenen Beziehungen können wir diese auch in der Form schreiben

$$(2)\ i_1+2\,i_2+\dots+\alpha\, i_\alpha+k_1+2\,k_2+\dots+\beta\, k_\beta=\alpha\,\beta.$$

§ 80. Discriminanten.

Wir haben früher (§ 65) bewiesen, dafs sich eine ganze Funktion F(x) dann und nur dann in mehrfache Faktoren zerlegen läfst, wenn sie mit ihrer Ableitung F(x) einen Faktor gemein hat. Aus der Resultante von F(x) und F'(x), die in diesem Falle verschwinden mufs, läfst sich der Koeffizient a_0 der höchsten Potenz der Variabeln als Faktor herausheben, und man nennt dann den andern die **Discriminante** von F(x), bezeichnet ihn kurz durch $\mathbf{D}(F(x))$, so dafs also

$$\mathbf{R}(Fx), F'(x)) = a_0 \mathbf{D}(F(x))$$

gesetzt wird.

Aus dem Multiplikationstheorem für die Resultanten ergiebt sich leicht ein ebensolches für die Discriminanten in der Form

$$\mathbf{D}(AB) = (-1)^{\alpha\beta} \mathbf{D}(A)\mathbf{D}(B)\mathbf{R}^2(A, B)$$

aus der folgenden Umformung

$$\mathbf{D}(AB) = \mathbf{R}(AB, AB' + A'B)$$
$$= \mathbf{R}(A, AB' + A'B)\mathbf{R}(B, AB' + A'B) = \mathbf{R}(A, A'B)\mathbf{R}(B, AB')$$
$$= \mathbf{R}(A, A')\mathbf{R}(A, B)\mathbf{R}(B, A)\mathbf{R}(B, B')$$
$$= (-1)^{\alpha\beta}\mathbf{R}^2(A, B)\mathbf{D}(A)\mathbf{D}(B).$$

Das Multiplikationstheorem läfst sich auch leicht auf mehrere Faktoren ausdehnen in der Form

$$\mathbf{D}(A_1 A_2 \dots A_n) = (-1)^{\sum\limits_{i,k} \alpha_i \alpha_k} \prod_{i,k} \mathbf{R}^2(A_i, A_k) \prod_i \mathbf{D}(A_i),$$
$$(i, k = 1, 2, \dots n;\ i \neq k)$$

wenn mit $\alpha_1, \alpha_2, \dots \alpha_n$ die Grade von $A_1, A_2, \dots A_n$ bezeichnet werden. Den sehr einfachen Beweis hierfür überlassen wir dem Leser zu führen, und wollen nur noch bemerken, dafs sich hieraus, wenn F(x) in Linearfaktoren zerlegt ist in der Form

$$F(x) = a_0(x - a_1)(x - a_2)\dots(x - a_n),$$

als eine Darstellung der Discriminante der Ausdruck

$$\mathbf{D}(F(x)) = (-1)^{\frac{n(n-1)}{2}} a_0^{2n-2} \prod_{i,k}(a_i - a_k)^2$$
$$(i, k = 1, 2, \dots n;\ i \neq k)$$

ergiebt. Dieser wird häufig als Ausgangspunkt für die Discriminanten gewählt, wenn man die Theorie der symmetrischen Funktionen voraussetzen will (§ 100).

Für Funktionen zweiten, dritten und vierten Grades lassen sich die Discriminanten leicht bilden. Sie sind in folgender Tafel enthalten:

$F(x) = a_0 x^2 + a_1 x + a_2$	$D(F) = a_1^2 - 4 a_0 a_2$
$F(x) = a_0 x^3 + a_1 x^2 + a_2 x + a_3$	$D(F) = a_1^2 a_2^2 + 18 a_0 a_1 a_2 a_3 - 4 a_0 a_2^3 - 4 a_1^3 a_3 - 27 a_0^2 a_3^2$
$F(x) = a_0 x^4 + a_1 x^3 + a_2 x^2 + a_3 x + a_4$	$27 D(F) = 4 A^3 - B^2$ $A = a_2^2 - 3 a_1 a_3 + 12 a_0 a_4$ $B = 27 a_1^2 a_4 + 27 a_0 a_3^2 + 2 a_2^3 - 72 a_0 a_2 a_4 + 9 a_1 a_2 a_3$

§ 81. Elimination.

Sind $A(x, y)$ und $B(x, y)$ zwei ganze Funktionen der Unbestimmten x und y, und handelt es sich darum, für x und y solche Werte zu ermitteln, für die beide Funktionen gleichzeitig verschwinden, so kann die Lösung dieser Aufgabe in folgender Weise reduziert werden. Haben zunächst beide Funktionen $A(x, y)$ und $B(x, y)$ einen von den Veränderlichen abhängigen gröfsten gemeinschaftlichen Teiler $T(x, y)$, so gehören diejenigen Werte von x, y, die diesen zum Verschwinden bringen, zu den gesuchten Werten. Alle andern müssen den beiden Gleichungen genügen, die man erhält, wenn man den gemeinschaftlichen Teiler aus den Funktionen entfernt.

Die Aufgabe, zwei solche teilerfremde Funktionen $A(x, y)$ und $B(x, y)$ durch geeignete Wertepaare zum Verschwinden zu bringen, läfst sich mit Hülfe der Resultanten beträchtlich vereinfachen. Fassen wir nämlich beide Funktionen als

Dies läfst sich auf folgendem Wege zeigen. Nach § 77 I, § 78 III ist

$$\begin{aligned} \mathbf{R}(F(x),\ F(x+u)) &= \mathbf{R}(F(x-u),\ F(x)) \\ &= (-1)^n\, \mathbf{R}(F(x),\ F(x-u)). \end{aligned}$$

Da aber

$$F(x \pm u) = F(x) \pm u\, F_1(x, \pm u)$$

ist, so folgt aus § 77 III, IV

$$\begin{aligned} \mathbf{R}(F(x),\ F(x \pm u)) &= \mathbf{R}(F(x),\ \pm u\, F_1(x, \pm u)) \\ &= (\pm u)^n\, \mathbf{R}(F(x),\ F_1(x, \pm u)), \end{aligned}$$

und daraus ergiebt sich dann

$$\mathbf{R}(F(x),\ F_1(x, u)) = \mathbf{R}(F(x),\ F_1(x, -u)),$$

oder dafs

$$G(u) = G(-u)$$

ist, d. h. $G(u)$ ist eine Funktion, die nur gerade Potenzen von u enthält. Da $F_1(x, u)$ für $u = 0$ in $F'(x)$ übergeht, so ist das von u unabhängige Glied in $G(u)$ gleich

$$\mathbf{R}(F(x),\ F'(x)) = a_0\, \mathbf{D}(F(x)),$$

also von Null verschieden. Was endlich den Grad von $G(u)$ betrifft, so beachten wir, dafs $F_1(x, u)$ die Form

$$F_1(x, u) = b_0 x^{n-1} + b_1 x^{n-2} + \dots + b_{n-2} x + b_{n-1},$$

wo die in $b_0, b_1, \dots b_{n-1}$ auftretenden Glieder der höchsten Potenz von u die folgenden sind:

$$n_1 a_0,\ n_2 a_0 u,\ n_3 a_0 u^2, \dots n_1 a_0 u^{n-2}, a_0 u^{n-1}.$$

Der in der Determinante $\mathbf{R}(F(x), F_1(x, u))$ auftretende Koeffizient mit der höchsten Potenz von u ist enthalten in $a_0^{n-1} b_{n-1}^n$, also, abgesehen von einem Zahlenkoeffizienten, gleich $a_0^{n-1} a_0^n = a_0^{2n-1}$, und die Potenz ist die $n(n-1)$te.

X. Abschnitt.

Wurzeln algebraischer Gleichungen.

---

§ 83. Irrationale und komplexe Wurzeln.

Ist F(x) eine ganze Funktion von x, so nennt man jeden Wert von x, der die Gleichung $F(x) = 0$ befriedigt, eine **Wurzel** dieser Gleichung (§ 59). Dafs es nur eine beschränkte Anzahl von solchen geben kann, nämlich höchstens soviel als der Grad von F(x) angiebt, wissen wir schon aus früheren Untersuchungen (§ 62). Wir haben auch schon Methoden kennen gelernt, mit deren Hülfe man sie bestimmen kann, wenn sie rationale Zahlen sind (§ 59). Solche rationalen Wurzeln existieren aber nur in speziellen Fällen, und wir stellen uns jetzt auf einen höheren Standpunkt, indem wir zunächst irrationale, dann aber auch komplexe Wurzeln mit in Betracht ziehen.

Eine irrationale reelle Wurzel wird immer so bestimmt, dafs man zwei rationale Zahlen angiebt, zwischen denen sie gelegen ist. Sind a und b zwei beliebige reelle Zahlen, von denen $a < b$ sei, so sagt man von allen Werten einer Variablen x, die der Bedingung

$$a < x < b$$

genügen, in Anlehnung an eine geometrische Ausdrucksweise, dafs sie im Intervalle $a \ldots b$ liegen; a nennt man die untere, b die obere Grenze des Intervalles. Liegt ein Wert x_0 im Intervall $a \ldots b$, so nennt man dieses auch ein Intervall um den Wert x_0. Wird eine irrationale Wurzel hiernach durch die Angabe eines Intervalles um sie bestimmt, so mufs es möglich sein, dieses Intervall beliebig einzuengen.

Man ist imstande dies auszuführen, wenn man allgemein die **Anzahl der reellen Wurzeln einer Gleichung in einem beliebigen Intervall** bestimmen kann. Denn dann kann man es in kleinere Teile zerlegen, zunächst so, dafs die einzelnen Wurzeln isoliert werden, und sodann jedes Intervall, in dem eine Wurzel gelegen ist, durch weitere Teilung, z. B. fortgesetzte Halbierung, so klein machen, dafs seine Gröfse unter jeden beliebigen Wert herabsinkt.

Wenn man auch **komplexe Wurzeln** in Betracht zieht, wie das in der Folge geschehen wird, sind von dieser die reellen und imaginären Bestandteile in ähnlicher Weise zu bestimmen. Um deutlich hervortreten zu lassen, dafs der Variabeln auch komplexe Werte beigelegt werden sollen, soll diese fortan durch den Buchstaben z bezeichnet werden, ihr reeller Bestandteil mit x, der imaginäre mit y, so dafs also

$$z = x + iy$$

gesetzt ist. Ist nun $F(z)$ eine ganze rationale Funktion dieser Variablen, so läfst sie sich stets in der Form darstellen

$$F(z) = X(x, y) + i\,Y(x, y),$$

wobei dann $X(x, y)$ und $Y(x, y)$ ganze Funktionen der beiden reellen Variabeln x und y mit reellen Koeffizienten sind.

Wenn $F(z) = 0$ wird, so mufs

$$X(x, y) = 0, \quad Y(x, y) = 0$$

sein, und es läfst sich daher die Theorie der **komplexen Wurzeln** als ein spezieller Fall der Theorie der **gemeinschaftlichen Wurzelsysteme zweier Gleichungen** mit zwei Unbekannten x und y auffassen.

Diese Theorie gewinnt bedeutend an Anschaulichkeit, wenn man ihr eine geometrische Einkleidung giebt und die Variabeln x und y als Koordinaten eines Punktes auffafst, bezogen auf ein rechtwinkliges Cartesisches Koordinatensystem. Die Gleichungen $X(x, y) = 0, Y(x, y) = 0$ stellen dann zwei Kurven dar, die den Gleichungen gemeinschaftlichen Wurzelsysteme entsprechen den Koordinaten der Schnittpunkte der beiden Kurven. Werden nun diese Schnittpunkte so bestimmt, dafs man für ihre Koordinaten Intervalle angiebt, in denen sie gelegen sind, so kommt dies geometrisch ausgedrückt, auf die Bestimmung eines kleinen rechteckigen

oder quadratischen Gebietes zurück, das die Schnittpunkte einschliefst. Allgemein kann man auch als Begrenzung des Gebietes jede geschlossene sich selbst nicht schneidende Kurve anwenden. Es ist aber von besonderem Vorteil, als solche den Kreis zu gebrauchen, und zwar aus folgendem Grunde. Ist z_0 ein beliebiger Wert von z, r eine positive Gröfse, so wird durch die Bedingung

$$|z - z_0| < r$$

ausgedrückt, dafs die der Variabeln z entsprechenden Punkte innerhalb eines solchen Kreises um den Punkt z_0 mit dem Radius r liegen. Setzt man nämlich

$$z = x + iy, \qquad z_0 = x + iy_0,$$

so genügen die reellen Variabeln x und y der Ungleichung

$$(x - x_0)^2 + (y - y_0)^2 < r^2.$$

Man ist imstande, alle komplexen Wurzeln zu bestimmen, wenn man allgemein feststellen kann, wie viele komplexe Wurzeln in einem beliebigen Bereiche gelegen sind. Denn dann kann man den Bereich zerteilen, die Wurzeln isolieren und weiterhin in ein beliebig kleines Gebiet einengen.

---

§ 84. Stetigkeit der ganzen Funktionen. Verhalten „im Unendlichen“.

I. Um jeden Wert z_0 einer komplexen Variabeln z kann man einen Bereich angeben, in dem der absolute Betrag der Wertänderung einer ganzen Funktion F(z), die Gröfse $|F(z) - F(z_0)|$, kleiner ist als eine beliebige positive Gröfse, wie klein diese auch sei.

Ausführlicher ausgesprochen heifst dies: Wenn δ eine beliebige noch so kleine positive Gröfse ist, so kann man für jeden Wert von z_0 immer eine positive Gröfse r bestimmen, dafs für alle Werte von z, die der Bedingung $|z - z_0| < r$ genügen, $|F(z) - F(z_0)| < \delta$ wird. Der Beweis für diese Möglichkeit läfst sich folgendermafsen erbringen. Nach der Taylorschen Entwickelung (§ 6) hat man

$$F(z) = F(z_0) + \sum_k \frac{F^{(k)}(z_0)}{k!}(z - z_0)^k \qquad k = 1, 2, \ldots n)$$

und daher

$$F(z) - F(z_0) = \sum_k \frac{F^{(k)}(z_0)}{k!}(z - z_0)^k$$

$$< \sum_k \left| \frac{F^{(k)}(z_0)}{k!} \right| . \left| z - z_0 \right|^k.$$

Nimmt man nun eine positive Gröfse g, die die Beträge

$$|F'(z_0)|, \left| \frac{F''(z_0)}{2!} \right|, \ldots \left| \frac{F^{(n)}(z_0)}{n!} \right|$$

übertrifft, so folgt weiter

$$|F(z) - F(z_0)| < g \sum_k |z - z_0|^k = g\, |z - z_0| \frac{1 - |z - z_0|^n}{1 - |z - z_0|}$$

Beschränkt man nun z vorläufig auf den Bereich $|z - z_0| < 1$, so ist für diesen

$$|F(z) - F(z_0)| < g \frac{|z - z_0|}{1 - |z - z_0|}.$$

Damit nun die Gröfse $g \frac{|z - z_0|}{1 - |z - z_0|} < \delta$ wird, hat man den Bereich weiter so einzuschränken, dafs

$$|z - z_0| < \frac{\delta}{g + \delta}$$

ist. Wählt man also $r \leqq \frac{\delta}{g + \delta}$, so ist $|F(x) - F(z_0)| < \delta$.

Man pflegt diesen Satz gewöhnlich kurz so auszusprechen: Jede ganze Funktion ist für jeden Wert der Variabeln stetig.

II. Man kann für die Variable z einen solchen Bereich bestimmen, dafs aufserhalb desselben der absolute Betrag der Differenz des Verhältnisses einer ganzen Funktion F(z) zu ihrem höchsten Gliede $a_0 z^n$ und der positiven Einheit kleiner ist als eine beliebige positive Gröfse, wie klein diese auch gewählt sei.

Ist also δ eine beliebig kleine positive Gröfse, so kann man immer eine positive Gröfse r so bestimmen, dafs für alle Werte von z, die der Bedingung $|z| > r$ genügen, der Ausdruck

$$\left|\frac{F(x)}{a_0 z^n} - 1\right| < \delta$$

ist. Den Beweis kann man mit Hülfe des Satzes I führen, indem man

$$z = \frac{1}{z'}, \qquad \frac{F(z)}{z^n} = F'(z')$$

setzt, wo dann $F'(z')$ eine ganze Funktion der Variabeln z' ist. Diese ist für jeden Wert von z' stetig, also auch für den Wert $z' = 0$, wofür sie den Wert $F'(0) = a_0$ annimmt. Daher kann man eine positive Gröfse r' so bestimmen, dafs für $|z'| < r'$ stets $|F'(z') - a_0|$ beliebig klein, also kleiner als $a_0 \delta$ ist. Dann ist aber $|z| > \frac{1}{r'}$, und

$$\left|\frac{F(z)}{a_0 z^n} - 1\right| < \delta.$$

Hat man also r' auf die angegebene Weise bestimmt, so braucht man nur $r \geqq \frac{1}{r'}$ anzunehmen.

III. Man kann einen solchen Bereich der Variabeln z bestimmen, dafs aufserhalb desselben der absolute Betrag einer ganzen Funktion F(z) gröfser als eine beliebige positive Gröfse ist, wie grofs diese auch gewählt sein mag.

Zerlegen wir nämlich F(z) in folgender Weise

$$F(z) = a_0 z^n \left\{1 + \left[\frac{F(z)}{a_0 z^n} - 1\right]\right\},$$

so folgt hieraus

$$|F(z)| > |a_0 z^n| \cdot \left[1 - \left|\frac{F(z)}{a_0 z^n} - 1\right|\right].$$

Ist nun α eine beliebige positive Gröfse, die kleiner als 1 ist, so bestimmen wir zunächst einen Bereich, aufserhalb

dessen $\left|\frac{F(z)}{a_0 z^n} - 1\right| < \alpha$ ist, und dann einen Bereich, aufserhalb dessen

$$|z^n| > \frac{\omega}{|a_0|(1-\alpha)}$$

ist. Für den gemeinsamen Aufsenbereich ist denn $|F(z)| > \omega$, und hierin kann ω jede positive Gröfse bedeuten.

Dieser Satz ist insofern von Wichtigkeit, als er betreffs der Wurzel eine, wenn auch nur rohe Grenze, bestimmen lehrt, unterhalb deren die absoluten Beträge der sämtlichen Wurzeln liegen. Das Problem der Bestimmung aller Wurzeln überhaupt ist damit auf das Problem der Bestimmung aller Wurzeln in einem Gebiet oder Intervall reduziert.

---

§ 85. Fundamentalsätze über die Existenz reeller Wurzeln.

Aus der Stetigkeit der ganzen Funktionen ergiebt sich nun sofort folgender Fundamentalsatz über die Existenz einer reellen Wurzel:

I. Wenn eine ganze Funktion an den Endpunkten eines Intervalles reelle Werte von verschiedenem Vorzeichen hat, so nimmt sie im Innern des Intervalles mindestens für einen Wert der Variabeln den Wert Null an.

Denn sonst würde sie im Intervall einmal plötzlich von einem positiven zu einem negativen Wert überspringen müssen, was durch ihre Stetigkeit ausgeschlossen ist.

Aus diesem Satze lassen sich mit Berücksichtigung des Satzes III des vorigen Paragraphen einige Folgerungen ziehen. Man kann nämlich der reellen Variabeln einen so grofsen Wert erteilen, dafs das Vorzeichen von $F(x)$ mit dem seines höchsten Gliedes $a_0 x^n$ übereinstimmt, und dabei das Vorzeichen von x sowohl positiv als auch negativ annehmen. Ist nun der Grad n von $F(x)$ eine ungerade Zahl, so nimmt $a_0 x^n$ entgegengesetzte Zeichen an, wenn das Zeichen von x geändert wird, weil $x^n = -(-x)^n$ ist. Jede Gleichung von ungerader Gradzahl hat demnach mindestens eine reelle Wurzel. Beachtet man ferner,

dafs das konstante Glied a_n von $F(x)$ gleich $F(0)$ ist, und dafs man x einen so hohen positiven Wert geben kann, dafs das Vorzeichen von $F(x)$ mit dem von a_0 übereinstimmt, so erkennt man: *Haben die Koeffizienten der höchsten und niedrigsten Potenz entgegengesetzte Vorzeichen, so hat die Gleichung mindestens eine positive Wurzel.* Ähnlich so ergiebt sich: *Haben der Koeffizient der höchsten und niedrigsten Potenz dasselbe Vorzeichen und ist die Gradzahl ungerade, so besitzt die Gleichung mindestens eine negative Wurzel.* Ist a eine positive Zahl, so hat die Gleichung $F(x) = x^n - a$ stets eine positive Wurzel, und wenn n gerade ist, auch noch eine negative; dafs keine weiteren reellen Wurzeln vorhanden sind, läfst sich leicht mit Berückstigung des Umstandes ableiten, dafs $|x|^n$ gleichzeitig mit dem absoluten Wert $|x|$ wächst.

Lehrt uns der Satz I die Existenz einer reellen Wurzel in einem bestimmten Falle kennen, so kann er doch nicht als ein allgemeines Kriterium für die Existenz einer Wurzel dienen, weil eine Funktion sehr wohl in einem Intervall verschwinden kann, ohne dafs ein Vorzeichenwechsel eintritt. Demgegenüber ist folgender Satz von Wichtigkeit:

II. *Wenn eine ganze Funktion $F(x)$, die zu ihrer Derivierten $F(x)$ teilerfremd ist, verschwindet für einen Wert, so kann man diesen in ein Intervall einschliefsen, in dem $F(x)$ das Vorzeichen einmal und nur einmal wechselt.*

Ist x_0 die Wurzel, also $F(x_0) = 0$, so mufs $F'(x_0) \neq 0$ sein, weil sonst $F(x)$ und $F'(x)$ den Faktor $(x - x_0)$ gemeinsein hätten. Benutzen wir nun die Entwickelung

$$F(x) = F(x_0) + \sum_k \frac{F^{(k)}(x_0)}{k!}(x - x_0)^k, \quad (k = 1, 2, \ldots n)$$

so folgt aus ihr, dafs

$$\frac{F(x)}{x - x_0} - F'(x_0) = (x - x_0) \sum_k \frac{F^{(k)}(x_0)}{k!}(x - x_0)^{k-1}$$
$$(k = 2, 3, \ldots n)$$

als eine ganze Funktion von $(x - x_0)$ angesehen werden kann, die für $x - x_0 = 0$ verschwindet. Daher kann man

um x_0 ein Intervall bestimmen, für dessen Werte x der Ausdruck

$$\left|\frac{F(x)}{x - x_0} - F'(x_0)\right| < |F'(x_0)|$$

ist, und also $\frac{F(x)}{x - x_0}$ dasselbe Vorzeichen hat wie $F'(x_0)$. Da nun aber $(x - x_0)$ in dem Intervall sein Vorzeichen einmal und nur einmal wechselt, so gilt dasselbe für $F(x)$.

Mit Hülfe dieses Satzes gelangen wir nun über die Anzahl der reellen Wurzeln in einem beliebigen Intervall zu folgendem Resultat:

III. Ist die ganze Funktion $F(x)$ zu ihrer Ableitung $F'(x)$ teilerfremd, so ist die Anzahl der reellen Wurzeln der Gleichung $F(x) = 0$ in einem beliebigen Intervalle gleich der Anzahl der Zeichenwechsel, die $F(x)$ erleidet, wenn die Variable x das Intervall stetig durchläuft.

Die hierbei auftretenden Zeichenwechsel bestehen abwechselnd aus Übergängen vom Positiven ins Negative und vom Negativen ins Positive. Es läſst sich aber auch ein Kriterium aufstellen, bei dem nur die eine Art von Übergängen in Betracht zu ziehen ist. Kehren wir noch einmal zu dem Beweise des Satzes II zurück und beachten wir, daſs wegen der Stetigkeit von $F'(x)$ sich um x_0 ein Intervall bestimmen läſst, für dessen Werte x die Funktion $F'(x)$ dasselbe Vorzeichen hat wie $F'(x_0)$, so erkennen wir, daſs es auch ein Intervall um x_0 giebt, in dem $\frac{F(x)}{F'(x)}$ dasselbe Vorzeichen hat wie $(x - x_0)$. Läſst man nun x wachsend das Intervall durchlaufen, so geht aber $(x - x_0)$ vom Negativen ins Positive über, also auch $\frac{F(x)}{F'(x)}$. Da das für jede reelle Wurzel gilt, so gelangen wir zu folgendem Satze:

IV. Sind $F(x)$ und $F'(x)$ zu einander teilerfremd, so stimmt die Anzahl der reellen Wurzeln der Gleichung $F(x) = 0$ in einem beliebigen Intervalle überein mit der Anzahl der Übergänge vom Negativen ins Positive, die das Verhältnis $F(x) : F'(x)$ beim Verschwinden erleidet, wenn x wachsend das ganze Intervall stetig durchläuft.

Es ist wohl zu beachten, dafs das Verhältnis $F(x) : F'(x)$ auch sonst noch Vorzeichenwechsel erleidet, nämlich wenn $F'(x)$ verschwindet, wobei dann immer ein Übergang (durchs Unendliche) vom Positiven ins Negative erfolgt. Hieraus folgt dann mit Zuhülfenahme des Satzes I, dafs zwischen je zwei aufeinanderfolgenden Wurzeln der Gleichung $F(x) = 0$ mindestens eine Wurzel von $F'(x) = 0$ liegt (und wenn mehrere, so eine ungerade Anzahl).

Wir haben früher (§ 65) gesehen, dafs man eine ganze Funktion auf rationalem Wege in ein Produkt von solchen Faktoren zerlegen kann, von denen jeder zu seiner Ableitung teilerfremd ist. Daraus ergiebt sich, dafs die entwickelten Sätze über die Existenz von reellen Wurzeln als wirkliche Kriterien für die reellen Wurzeln betrachtet werden können. Wir wollen dies ausdrücken durch den Satz:

V. Um jede Wurzel einer Gleichung $F(x) = 0$ kann man ein Intervall bestimmen, in dem ein rationaler Teiler von $F(x)$ sein Vorzeichen einmal und nur einmal wechselt.

---

§ 86. Einfachste Methode zur Bestimmung der reellen Wurzeln.

Die in dem vorigen Paragraphen entwickelten Kriterien für die Existenz reeller Wurzeln, die in den Sätzen II und III ihren Ausdruck finden, sind nun ohne Weiteres nicht anwendbar, so lange man kein Mittel hat, aus der Bestimmung des Vorzeichens einer Funktion für eine endliche Anzahl von Werten in einem Intervalle einen Schlufs auf den Vorzeichenverlauf in dem ganzen Intervalle zu ziehen.

Der Satz III wird nun anwendbar, wenn man ein Mittel angiebt, um das Intervall d zu bestimmen, in dem $F(x)$ sein Vorzeichen entweder beibehält oder nur ein einziges Mal wechselt. Dieses kritische Intervall braucht nur kleiner angenommen zu werden, als der Betrag der Differenz zweier Wurzeln. Wie wir nun früher gesehen haben, ist es stets möglich, eine Gleichung aufzustellen, der alle Differenzen genügen (§ 82). Sie lautet, wenn

$$F(x+u) = F(x) + u\,F_1(x, u)$$

gesetzt wird,

$$R[F(x),\ F_1(x, u)] = 0,$$

enthält nur Potenzen von u^2, und ihr konstantes Glied ist von 0 verschieden. Daher kann man stets eine untere Grenze für den absoluten Betrag der Wurzeln u bestimmen und diese dann als kritisches Intervall benutzen, in dem F(x) entweder sein Vorzeichen beibehalten mufs oder nur ein einziges Mal wechseln kann.

Obgleich sich an dieser Methode noch Vereinfachungen anbringen lassen, so dafs es gar nicht nötig ist, die Gleichung für u selbst erst aufzustellen, so ist sie doch aufserordentlich unbequem für die Anwendung, weil die Gröfse d einen sehr kleinen Wert hat, und man daher sehr viele Funktionswerte von F(x) zu berechnen hat, um über die Wurzelverteilung ein klares Bild zu gewinnen.

Demgegenüber ist es von aufserordentlicher Wichtigkeit, dafs der Satz IV des vorigen Paragraphen auf eine weit einfacher zu handhabende Methode führt, die zuerst von Sturm gefunden wurde, und zu deren Darlegung wir jetzt übergehen wollen.

---

§ 87. Excefs eines Funktionenverhältnisses in einem Intervall.

Durch die Untersuchungen in den vorhergehenden Paragraphen ist die Bestimmung der reellen Wurzeln einer Gleichung mit reellen Koeffizienten auf ein Problem zurückgeführt worden, das man in verallgemeinerter Form so aussprechen kann:

Es sind zwei Funktionen A(x) und B(x) gegeben, die nicht für gleiche Werte verschwinden, und es soll die Zahl bestimmt werden, die angiebt, wie oft das Verhältnis A(x):B(x) beim Verschwinden mehr vom Positiven ins Negative als vom Negativen ins Positive übergeht, wenn die Variable x stetig alle Werte eines Intervalles von a bis b durchläuft. Diese Zahl nennt man den Excefs des Funktionen-

verhältnisses $A(x):B(x)$ im Intervall $a \dots b$, und wir wollen sie kurz mit

$$\mathop{E}_{a}^{b}[A(x):B(x)]$$

bezeichnen.

Hiernach wird die Anzahl der reellen Wurzeln der Gleichung $F(x)=0$ in dem Intervalle $a \dots b$, wenn $F(x)$ und $F'(x)$ teilerfremd sind, einfach ausgedrückt durch

$$-\mathop{E}_{a}^{b}[F(x):F'(x)],$$

wobei $b > a$ angenommen ist.

Wir stellen nun über die Excesse eine Reihe von Gleichungen auf.

I. Es ist

$$\mathop{E}_{a}^{b}[A(x):B(x)] = -\mathop{E}_{b}^{a}[A(x):B(x)]$$

$$\mathop{E}_{a}^{b}[A(x):B(x)] + \mathop{E}_{b}^{c}[A(x):B(x)] = \mathop{E}_{a}^{c}[A(x):B(x)].$$

Die Richtigkeit der ersten Gleichung sieht man sofort ein, wenn man bedenkt, dafs bei einer entgegengesetzten Durchlaufungsrichtung die beiden Arten der Übergange vom Positiven ins Negative und vom Negativen ins Positive sich wechselseitig entsprechen. Die zweite Gleichung ist zunächst für den Fall unmittelbar einleuchtend, dafs b im Intervall $a \dots c$ liegt; dafs aber auch dann ihre Giltigkeit nicht aufhört, wenn b aufserhalb des Intervalles liegt, ergiebt sich leicht mit Zuhülfenahme der ersten Gleichung. Man kann die zweite Gleichung auch in der Form

$$\mathop{E}_{a}^{b}[A(x):B(x)] + \mathop{E}_{b}^{c}[A(x):B(x)] + \mathop{E}_{c}^{a}[A(x):B(x)] = 0$$

schreiben.

II. Haben $B(x)$ und $B_1(x)$ immer dasselbe Vorzeichen, wenn $A(x)$ verschwindet, so ist

$$\mathop{E}_{a}^{b}[A(x):B(x)] = \mathop{E}_{a}^{b}[A(x):B_1(x)].$$

Das trifft insbesondere zu, wenn

$$B(x) \equiv B_1(x) \mod A(x)$$

ist.

III. Bedeutet $C(x)$ eine Funktion, die im Intervall $a \ldots b$ des Excesses nicht verschwindet, so ist

$$\mathop{E}_a^b [C(x) : A(x)] = 0$$

$$\mathop{E}_a^b [A(x)\, C(x) : B(x)] = \mathop{E}_a^b [A(x) : B(x)\, C(x)] = \varepsilon \mathop{E}_a^b [A(x) : B(x)],$$

wo ε in der letztern Gleichung das konstante Vorzeichen bedeutet, das $C(x)$ im Intervall hat. Auch dies ist unmittelbar klar. Insbesondere aber folgt hieraus

$$\mathop{E}_a^b [A(x) : B(x)] = - \mathop{E}_a^b [- A(x) : B(x)]$$

$$= - \mathop{E}_a^b [A(x) : - B(x)] = \mathop{E}_a^b [- A(x) : - B(x)]$$

und bei zweimaliger Anwendung des Satzes

$$\mathop{E}_a^b [A(x)\, C(x) : B(x)\, C(x)] = \mathop{E}_a^b [A(x) : B(x)].$$

IV. Sind $A(x)$ und $B(x)$ Funktionen, die im Intervall $a \ldots b$ nicht gleichzeitig verschwinden, so ist

$$\mathop{E}_a^b [A(x)\, B(x) : C(x)] = \mathop{E}_a^b [A(x) : B(x)\, C(x)]$$

$$+ \mathop{E}_a^b [B(x) : A(x)\, C(x)].$$

Die Werte, für die $A(x)\, B(x)$ verschwindet, zerfallen dann nämlich in solche, für die $A(x) = 0$, $B(x) \neq 0$, und solche, für die $A(x) \neq 0$, $B(x) = 0$ ist; für erstere macht das Verhältnis $A(x)\, B(x) : C(x)$ dieselbe Vorzeichenänderung durch wie $A(x) : B(x)\, C(x)$, weil $B(x)$ innerhalb eines gewissen Intervalles konstantes Vorzeichen beibehält; für letztere erleidet das Verhältnis dieselbe Vorzeichenänderung wie $B(x) : A(x)\, C(x)$.

V. Vertauschen die beiden Funktionen eines Excesses ihre Rolle, so gilt der Satz:

$$\overset{b}{\underset{a}{E}}[A(x):B(x)] + \overset{b}{\underset{a}{E}}[B(x):A(x)]$$
$$= V[A(b),\ B(b)] - V[A(a),\ B(a)].$$

Hierbei bedienen wir uns, wie in der Folge des Zeichens V, um damit die Anzahl der Vorzeichenwechsel auszudrücken, die die in Klammern beigesetzten Gröfsen darbieten. Stehen also wie hier hinter V nur zwei Gröfsen, so hat V den Wert 0 oder 1, je nachdem diese beiden von gleichem oder entgegengesetztem Vorzeichen sind.

Der Ausdruck auf der linken Seite stellt die algebraische Summe der sämtlichen Vorzeichenwechsel dar, die das Verhältnis $A(x):B(x)$ erleidet, während die Variable x das Intervall von a bis b durchläuft, wenn man einen Übergang vom Positiven ins Negative als einen positiven, vom Negativen ins Positive als einen negativen betrachtet. Da immer zwei solche entgegengesetzte Vorzeichenwechsel aufeinander folgen, so kann die linke Seite nur die Werte 0, 1 und -1 annehmen. Im ersteren Falle hat das Verhältnis $A(a):A(a)$ dasselbe Vorzeichen wie $A(b):B(b)$, und demnach ist auch die rechte Seite gleich 0. Im zweiten Falle ist das Verhältnis $A(a):B(a) > 0$, dagegen $A(b):B(b)$ negativ und daher die rechte Seite gleich $+1$. Im dritten Falle endlich ist $A(a):B(a) < 0$ und $A(b):B(b) > 0$, die rechte Seite also gleich -1.

VI. Die zuletzt abgeleitete Gleichung ist vollständig symmetrisch in Bezug auf $A(x)$ und $B(x)$ gebaut. Deshalb können wir, ohne die Allgemeinheit zu beeinträchtigen, annehmen, dafs der Grad α von $A(x)$ den Grad β von $B(x)$ übertrifft. Dann kann man eine Funktion $A_1(x)$ vom Grade γ so bestimmen, dafs

$$A(x) \equiv A_1(x) \bmod B(x),\quad \gamma < \beta$$

ist, und dann den Satz II in Anwendung bringen in der Form

$$\overset{b}{\underset{a}{E}}[B(x):A(x)] = \overset{b}{\underset{a}{E}}[B(x):A_1(x)].$$

Hieraus ergiebt sich eine Reduktionsformel

$$\overset{b}{\underset{a}{E}}[A(x):B(x)] + \overset{b}{\underset{a}{E}}[B(x):A_1(x)] = V[A(b), B(b)] - V[A(a), B(a)],$$

durch die die Bestimmung des Excesses von $A(x):B(x)$ auf die von $B(x):A_1(x)$ zurückgeführt wird, wobei die neuen Funktionen von niederem Grade sind. Diese Bemerkung ist von grofser Wichtigkeit, da sie zur Berechnung eines Excesses ausgenutzt werden kann.

Ersetzt man $A_1(x)$ durch $-C(x)$, so verändert sich die Gestalt der Formel nur wenig. Wir erhalten nach III

$$\overset{b}{\underset{a}{E}}[A(x):B(x)] = V[A(b), B(b)] - V[A(a), B(a)] + \overset{b}{\underset{a}{E}}[B(x):C(x)].$$

Übrigens ist hierbei zu bemerken, dafs die Giltigkeit dieser Gleichung keineswegs an die Kongruenz

$$A(x) \equiv -C(x) \mod B(x)$$

gebunden ist. Sie gilt überhaupt, wenn $A(x)$ und $C(x)$ entgegengesetztes Vorzeichen haben für jede Wurzel von $B(x)$ in dem Intervall $a \dots b$.

---

§ 88. Vorzeichenwechsel. Harriotscher Satz.

Die letzten Gleichungen des vorigen Paragraphen lassen erkennen, dafs ein Zusammenhang besteht zwischen dem Excefs eines Funktionenverhältnisses und den Vorzeichenwechseln von zwei reellen Gröfsen; die in der letzten Gleichung enthaltene Reduktionsmethode kann dazu benutzt werden, um den Excefs eines Funktionenverhältnisses allgemein zu bestimmen. Da hierbei eine Reihe von Gröfsen auftreten, deren Vorzeichenwechsel in Betracht zu ziehen sind, so wollen wir hierüber einige Bemerkungen voraufschicken.

Sind $a, b, c, \dots$ eine Reihe von reellen Gröfsen, so bezeichnen wir die Anzahl der Vorzeichenwechsel, die sich darbieten, wenn man immer zwei aufeinanderfolgende mit einander hinsichtlich ihres Vorzeichens vergleicht, kurz durch

das Zeichen $V(a, b, c, \ldots)$. Ein Zweifel kann nur auftreten, wenn eine der Gröſsen verschwindet, und wir wollen dann festsetzen, daſs das Vorzeichen mit dem der voraufgehenden Gröſse als gleich betrachtet werden soll. Es ist nun leicht, die Richtigkeit folgender Gleichungen zu erkennen:

$$\text{(I)} \qquad V(a_0, a_1, \ldots a_n) = \sum_i V(a_{i-1}, a_i) \qquad (i = 1, 2, \ldots n)$$

$$\text{(II)} \qquad V(a_0, a_1, \ldots a_k, a_{k+1}, \ldots a_n) = V(a_0, a_1, \ldots a_k) + V(a_k, a_{k+1}, \ldots a_n),$$

von denen die zweite übrigens aus der ersten abgeleitet werden kann.

Es ist einleuchtend, daſs man in den Symbol $V(a, b, c \ldots)$ immer ein Glied entfernen kann, das entweder mit seinem vorangehenden, oder mit seinem folgenden gleiches Vorzeichen besitzt. Ist dagegen das Vorzeichen des entfernten Gliedes entgegengesetzt dem Vorzeichen der beiden angrenzenden Glieder, so wird der Wert des Symbols V hierbei um die Zahl 2 vermindert. Man kann dies zum Ausdruck bringen durch die Gleichung

$$\text{(III)} \qquad V(a_0, a_1, \ldots a_{i-1}, a_i, a_{i+1}, \ldots a_n)$$
$$= V(a_1, \ldots a_{i-1}, a_{i+1} \ldots a_n) + 2\, V(a_{i-1}, a_i)\, V(a_i, a_{i+1}).$$

Insbesondere ist

$$V(\ldots a, b, \ldots) = V(\ldots a, a + b, b, \ldots)$$
$$V(a, -a) = V(a, b, -a).$$

Sind a und b von Null verschieden, so ist immer

$$V(a, b) + V(a, -b) = 1.$$

Da nun

$$V(a_0, a_1, \ldots a_n) = \sum_i V(a_{i-1}, a_i) \qquad (i = 1, \ldots n)$$
$$V(a_0, -a_1, \ldots (-1)^n a_n) = \sum_i V((-1)^{i-1} a_{i-1}, (-1)^i a_i)$$

ist, so ergiebt sich durch Addition

$$\text{(IV)} \qquad V(a_0, a_1, \ldots a_n) + V(a_0, -a_1, a_2, -a_3, \ldots (-1)^n a_n) = n.$$

V) Mit Hülfe dieser Betrachtungen ist man auch imstande, einen von Harriot aufgestellten Satz über die Anzahl der positiven Wurzeln herzuleiten, die eine Gleichung höchstens haben kann.

Betrachten wir zu diesem Zwecke die ganze Funktion

$$F(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n$$

und bilden das Produkt

$$F(x)(x - h) = a_0' x^{n+1} + a_1' x^n + \ldots + a'_{n+1},$$

so ist

$$a_0 = a_0', \quad a_i = a_i' + a_{i-1} h, \quad a'_{n+1} = - a_n h. \qquad (i = 1, \ldots n)$$

Wir formen nun den Ausdruck $V(a_0', a_1', \ldots a'_{n+1})$ um, wobei wir annehmen, dafs $h > 0$ sei. Ersetzen wir a_0' durch a_0, und schalten zwischen a_0 und a_1' die Gröfse a_1 ein, wie es gestattet ist, so ergiebt sich

$$V(a_0', a_1', \ldots a'_{n+1}) = V(a_0, a_1, a_1', \ldots, a'_{n+1})$$

und weiter, wenn man (III) anwendet,

$$\begin{aligned} &V(a_0', a_1', \ldots a'_{n+1}) \\ &= V(a_0, a_1, a_2', \ldots a'_{n+1}) + 2\, V(a_1', a_1)\, V(a_1', a_2'). \end{aligned}$$

Schaltet man nun weiter zwischen a_1 und a_2' den Koeffizienten a_2 ein, so erhält man

$$\begin{aligned} V(a_0, a_1, a_2', \ldots a'_{n+1}) &= V(a_0, a_1, a_2, a_2', \ldots a'_{n+1}) \\ = V(a_0, a_1, a_2, a_3', \ldots, a'_{n+1}) &+ 2\, V(a_2', a_2)\, V(a_2', a_3'). \end{aligned}$$

Allgemein ergiebt sich

$$\begin{aligned} &V(a_0, a_1, \ldots a_i, a'_{i+1}, \ldots a'_{n+1}) \\ &= V(a_0, a_1, \ldots a_i, a_{i+1}, a'_{i+1}, \ldots a'_{n+1}) \\ &= V(a_0, a_1, \ldots a_i, a_{i+1}, a'_{i+2}, \ldots a'_{n+1}) \\ &\quad + 2\, V(a'_{i+1}, a_{i+1})\, V(a'_{i+1}, a'_{i+2}) \end{aligned}$$

und zuletzt

$$\begin{aligned} V(a_0, a_1, \ldots a_n, a'_{n+1}) &= V(a_0, a_1, \ldots a_n) + V(a_n, a'_{n+1}) \\ &= V(a_0, a_1, \ldots a_n) + 1. \end{aligned}$$

Das Gesamtergebnis dieser Transformation kann man durch Addition zusammenfassen in die Formel

$$\begin{aligned} &V(a_0', a_1', \ldots a'_{n+1}) \\ &= 1 + V(a_0, a_1, \ldots a_n) + 2 \sum_i V(a_i', a_i)\, V(a_i', a'_{i+1}) \end{aligned}$$

$$(i = 1, 2, \ldots n)$$

und aussprechen durch den Satz:

Ist h eine positive Gröfse, so hat die Funktion $F(x)(x-h)$ eine ungerade Anzahl von Zeichenwechseln mehr in seinen Koeffizienten als $F(x)$.

Wendet man den Satz wiederholt an, so erkennt man:

Sind $h_1, h_2, \ldots h_r$ positive Gröfsen, so übertrifft die Anzahl der Zeichenwechsel der Koeffizienten der Funktion $F(x)(x-h_1)(x-h_2)\ldots(x-h_r)$ die Anzahl der Zeichenwechsel in den Koeffizienten von $F(x)$ mindestens um r, und der etwaige Überschufs ist stets eine gerade Zahl.

Hat nun $F(x)$ keine positive Wurzel, so bietet die Reihe der Koeffizienten eine gerade Anzahl von Zeichenwechseln dar (§ 85), und daraus endlich folgt der Satz von Harriot:

Die Anzahl der positiven Wurzeln einer Gleichung ist höchstens gleich der Anzahl der Zeichenwechsel in der Reihe der Koeffizienten, und die Differenz beider Anzahlen ist stets eine gerade Zahl.

---

§ 89. Bestimmung des Excesses eines Funktionenverhältnisses. Sturmscher Satz.

Die in den beiden vorhergehenden Paragraphen entwickelten Sätze wollen wir nun dazu benutzen, um die vollständige Berechnung eines Excesses $\mathop{E}\limits_a^b [F(x) : F_1(x)]$ zu zeigen. Wir nehmen an, dafs eine Reihe von Funktionen

$$F(x),\ F_1(x),\ F_2(x), \ldots F_n(x)$$

abgeleitet sei, die mit den beiden gegebenen Funktionen $F(x), F_1(x)$ beginnt, von denen die erstere die zweite im Grade übertreffen mag, mit einer im Intervall $a \ldots b$ ihr Vorzeichen beibehaltenden Funktion $F_n(x)$ endigt, und deren inneren Glieder die folgende Eigenschaft besitzen: Wenn $F_i(x)$ für einen Wert im Intervall $a \ldots b$ verschwindet, so haben die beiden benachbarten

$F_{i-1}(x)$ **und** $F_{i+1}(x)$ **entgegengesetzte Vorzeichen.** Dies trifft im besonderen zu, wenn

$$F_{i-1}(x) \equiv - F_{i+1}(x) \bmod F_i(x)$$

ist, so dafs die Möglichkeit der Bestimmung solcher Funktionen damit sichergestellt ist. Wir erhalten dann folgende Gleichungen

$$\mathop{E}_a^b [F(x) : F_1(x)] - \mathop{E}_a^b [F_1(x) : F_2(x)]$$
$$= V[F(b), F_1(b)] - V[F(a), F_1(a)]$$
$$\mathop{E}_a^b [F_1(x) : F_2(x)] - \mathop{E}_a^b [F_2(x) : F_3(x)]$$
$$= V[F_1(b), F_2(b)] - V[F_1(a), F_2(a)],$$

allgemein

$$\mathop{E}_a^b [F_{i-1}(x) : F_i(x)] - \mathop{E}_a^b [F_i(x) : F_{i+1}(x)]$$
$$= V[F_{i-1}(b), F_i(b)] - V[F_{i-1}(a), F_i(a)],$$
$$(i = 1, 2, \ldots n-1)$$

endlich aber, weil $F_n(x)$ von konstantem Vorzeichen und

$$\mathop{E}_a^b [F_n(x) : F_{n-1}(x)] = 0$$

ist,

$$\mathop{E}_a^b [F_{n-1}(x) : F_n(x)] = V[F_{n-1}(b), F_n(b)] - V[F_{n-1}(a), F_n(a)].$$

Durch Addition ergiebt sich aus diesen Gleichungen

$$\mathop{E}_a^b [F(x) : F_1(x)]$$
$$= V[F(b), F_1(b), \ldots F_n(b)] - V[F(a), F_1(a), \ldots F_n(a)].$$

Ist nun $F_1(x) = F'(x)$ die Ableitung von $F(x)$ und $a > b$, so **giebt die linke Seite der vorstehenden Gleichung die Anzahl der reellen Wurzeln der Gleichung** $F(x) = 0$ **zwischen den Grenzen** a **und** b **an.** Dieser **Satz** ist zuerst von **Sturm** aufgestellt worden. Nennt man die Reihe

$$F(x), F_1(x) = F'(x), F_2(x), \ldots F_n(x)$$

eine Sturmsche Kette, so kann man den Sturmschen Satz folgendermafsen formulieren:

Die Anzahl der reellen Wurzeln der Gleichung $F(x) = 0$ in einem beliebigen Intervalle ist gleich dem Überschuſs der Anzahl der Zeichenwechsel der Sturmschen Kette an der unteren Grenze über die Anzahl der Zeichenwechsel an der oberen Grenze des Intervalles.

---

§ 90. Anwendungen des Sturmschen Satzes.

Wir wollen jetzt an einigen Beispielen die Anwendung des Sturmschen Satzes erläutern. Mit V_a bezeichnen wir die Anzahl der Zeichenwechsel in der Sturmschen Kette für den Wert $x = a$, speziell mit $V_{+\infty}$, $V_{-\infty}$ für einen beliebig hohen positiven oder negativen Wert von x.

1) $F(x) = x^m - 1$.

Sturmsche Kette: $x^m - 1,\ m x^{m-1},\ +1$.

$$V_\infty = V(+1, +1, +1) = 0$$
$$V_0 = V(-1, 0, +1) = 1$$
$$V_{-\infty} = V((-1)^m, (-1)^{m-1}, +1) = 1 + \frac{1 + (-1)^m}{2}.$$

Man hat demnach für die Anzahl der reellen, der positiven und der negativen Wurzeln resp.

$$V_{-\infty} - V_{+\infty} = 1 + \frac{1 + (-1)^m}{2}$$
$$V_0 - V_{+\infty} = 1$$
$$V_{-\infty} - V_0 = \frac{1 + (-1)^m}{2}.$$

2) $F(x) = x^2 + ax + b$

Sturmsche Kette: $x^2 + ax + b,\ 2x + a,\ a^2 - 4b = D$.

$$V_{+\infty} = V(+1, +1, D) = V(1, D)$$
$$V_0 = V(b, a, D)$$
$$V_{-\infty} = V(+1, -1, D) = V(1, D) + 2\,V(-1, D)$$
$$= 1 + V(-1, D).$$

Aus

$$V_{-\infty} - V_{+\infty} = 2\,V(-1, D)$$

folgt, daſs für $D < 0$ überhaupt keine reellen Wurzeln vor-

handen sind. Ist aber $D > 0$, so ergeben sich zwei reelle Wurzeln, und zwar positive in der Anzahl $V(1, a, b)$, negative in der Anzahl $V(1, -a, b)$.

3) $F(x) = x^3 + a x^2 + b x + c$.

Sturmsche Kette: $x^3 + a x^2 + b x + c,\ 3 x^2 + 2 a x + b,$

$$2(a^2 - 3b)x + (ab - 9c),$$

$$a^2 b^2 + 18 abc - 4 b^3 - 4 a^3 c - 27 c^2 = D.$$

$$V_{+\infty} = V(+1, +1, a^2 - 3b, D) = V(1, a^2 - 3b, D).$$

$$V_0 = V(c, b, ab - 9c, D).$$

$$V_{-\infty} = V(-1, +1, -(a^2 - 3b), D)$$
$$= 1 + V(1, -(a^2 - 3b), D).$$

Man erhält, da $V(1, a^2 - 3b, D) + V(1, -(a^2 - 3b), D) = 2$ ist,

$$V_{-\infty} - V_{+\infty} = 3 - 2 V(1, a^2 - 3b, D).$$

Ist $D < 0$, so ist nur eine reelle Wurzel vorhanden. Bei $D > 0$ ergeben sich deren drei, zugleich aber zeigt sich, dafs dann $a^2 - 3b > 0$ sein mufs, weil sonst $V_{-\infty} - V_{+\infty}$ negativ würde. Daher ist $V_{+\infty} = 0$, $V_{-\infty} = 3$. Die Anzahl der positiven Wurzeln wird angegeben durch $V(c, b, ab - 9c, D)$ und daher die der negativen durch $V(-c, b, -(ab - 9c), D)$, weil die Summe dieser beiden Symbole (§ 88 IV) gleich 3 ist. Aus der Gleichung

$$9D = -4b(a^2 - 3b)^2 - 3(ab - 9c)^2$$
$$+ 4a(ab - 9c)(a^2 - 3b)$$

erkennt man aber, dafs für $D > 0$, wo auch $a^2 - 3b > 0$ ist, a und $ab - 9c$ dasselbe Vorzeichen haben, wenn $b > 0$ ist. Man erhält, wenn man $ab - 9c$ durch a ersetzt, was auch bei $b < 0$ gestattet ist, ein einfacheres Kriterium, demzufolge die Anzahl der positiven Wurzeln durch $V(1, a, b, c)$, die der negativen durch $V(1, -a, b, -c)$ ausgedrückt wird.

4) Aus der Theorie der trigonometrischen Funktionen lassen sich einige Beispiele für den Sturmschen Satz verwerten, dessen Anwendung dann auf algebraischem Wege zu Resultaten führt, die man sonst anders ableitet. Wir wollen hier nur ein solches Beispiel behandeln und schicken zu diesem Zwecke einige Bemerkungen voraus.

Die Lösungen der Gleichung $\cos n u = 0$ genügen alle der Kongruenz

$$n u \equiv \frac{\pi}{2} \mod \pi,$$

aus der

$$u \equiv \frac{\pi}{2 n} \mod \frac{\pi}{n}$$

folgt. Will man also alle Lösungen mod π erhalten, so hat man in den Ausdruck $\frac{\pi}{2 n} + k \frac{\pi}{n}$ für k die Werte $0, 1, \ldots n - 1$ einzusetzen, und diese liefern alle n verschiedenen Werte von $\cos u$. Nun ist es aber möglich, $\cos n u$ als rationale Funktion von $\cos u$ vom n ten Grade darzustellen. Es ist nämlich

$$\cos (k + 1) u + \cos (k - 1) u = 2 \cos u \cos k u.$$

Setzt man

$$x = \cos u, \qquad \varphi_k(x) = \cos k u,$$

so erhält man die Rekursionsformel

$$\varphi_{k+1}(x) - 2 x \varphi_k(x) + \varphi_{k-1}(x) = 0,$$

aus der nicht nur die Richtigkeit der Behauptung hervorgeht, sondern die auch zur Berechnung der Funktionen dienen kann, wenn man hinzunimmt, dafs

$$\varphi_1(x) = x, \quad \varphi_2(x) = 2 x^2 - 1$$

ist. Es ergiebt sich z. B.:

$$\varphi_3(x) = 4 x^3 - 3 x, \quad \varphi_4(x) = 8 x^4 - 8 x^2 + 1,$$
$$\varphi_5(x) = 16 x^5 - 20 x^3 + 5 x, \text{ u. s. w.}$$

Die Gleichung $\varphi_n(x) = 0$ mufs nun n reelle zwischen den Grenzen $+1$ und -1 gelegene Wurzeln besitzen, die den n verschiedenen Werten von $\cos u$ entsprechen. Dieses Resultat soll jetzt mit Hilfe des Sturmschen Satzes abgeleitet werden. Da allgemein

$$\varphi_{k+1}(x) \equiv - \varphi_{k-1}(x) \mod \varphi_k(x),$$

so haben die Funktionen

$$\varphi_n(x), \ \varphi_{n-1}(x), \ \varphi_{n-2}(x), \ldots \ \varphi_1(x), \ \varphi_0(x) = 1$$

die Eigenschaften wie bei einer Sturmschen Kette, und es folgt aus den Rekursionsformeln, dafs

$$\varphi_k(1) = 1, \qquad \varphi_k(-1) = (-1)^k$$

ist. Somit ist

$$V_{+1} = V(+1, +1, \ldots +1) = 0,$$
$$V_{-1} = V((-1)^n, (-1)^{n-1}, (-1)^{n-2}, \ldots +1) = n,$$
$$V_{-1} - V_{+1} = n,$$

also $E_{-1}^{+1}[\varphi_n(x) : \varphi_{n-1}(x)] = -n$. Daher geht $\frac{\varphi_n(x)}{\varphi_{n-1}(x)}$, das wegen seines Grades beim Verschwinden nur höchstens n Zeichenwechsel erleiden kann, wirklich n mal vom Negativen ins Positive über, während x das Intervall $-1 \ldots +1$ stetig durchläuft, hat also thatsächlich n Wurzeln.

Dieses Beispiel ist insofern sehr bemerkenswert, als wir hier gar nicht nötig haben, die Ableitung von $\varphi_n(x)$ zu betrachten, was allerdings erst durch den Erfolg gerechtfertigt wird. Will man dies von vornherein doch thun, so bietet sich folgender Weg dar. Aus der Gleichung $\cos nu = \varphi_n(x)$ ergiebt sich durch Differenzieren nach u:

$$n \sin nu = \varphi'_n(x) \sin u.$$

Beachtet man nun

$$\cos(n-1)u = \cos nu \cos u + \sin nu \sin u,$$

so erhält man

$$\varphi_{n-1}(x) = x\,\varphi_n(x) + \frac{1}{n}\varphi_n'(x)(1 - x^2)$$

und erkennt daraus sofort, dafs, wenn man sich auf das Intervall $-1 \ldots +1$ beschränkt, $\varphi_{n-1}(x)$ beim Verschwinden von $\varphi_n(x)$ dasselbe Vorzeichen hat wie $\varphi_n'(x)$ und es daher vollständig ersetzen kann.

---

§ 91. Excefs eines Funktionenverhältnisses auf geschlossener Bahn.

Wir verallgemeinern den früher erörterten Begriff des Excesses eines Funktionenverhältnisses in folgender Weise: Es seien A(x, y) und B(x, y) zwei ganze Funktionen, A und B die Kurven, die durch die Gleichungen

$$A(x, y) = 0, \qquad B(x, y) = 0$$

dargestellt werden. Nehmen wir nun einen Weg W hinzu, der von dem Punkte (x_0, y_0) zum Punkt (x_1, y_1) führt und durch keinen Schnittpunkt der beiden Kurven A und B hindurch geht, so soll der Ausdruck

$$\underset{(x_0, y_0)}{\overset{(x_1, y_1)}{E}}(w)\,[A(x, y) : B(x, y)]$$

angeben, wie oft das Verhältnis $A(x, y) : B(x, y)$ bei seinem Verschwinden mehr vom Positiven ins Negative, als vom Negativen ins Positive übergeht, wenn das Variabelnpaar (x, y) den Weg W von (x_0, y_0) stetig bis (x_1, y_1) durchläuft. Wir wollen nun annehmen, dafs dieser Weg eine in sich geschlossene Kurve S darstellt, die nicht nur ihrer Gestalt, sondern auch dem Durchlaufungsinne nach gegeben sein mufs, den Excefs in diesem Falle mit

$$\mathbf{E}_S\,[A(x, y) : B(x, y)]$$

bezeichnen und den Excefs des Funktionenverhältnisses $A(x, y) : B(x, y)$ auf der geschlossenen Bahn S nennen. Es ist dann stets

$$\text{(I)} \qquad \mathbf{E}_S\,[A(x, y) : B(x, y)] + \mathbf{E}_S\,[B(x, y) : A(x, y)] = 0.$$

Denn da das Verhältnis $A(x, y) : B(x, y)$ abwechselnd vom Positiven ins Negative und vom Negativen ins Positive übergeht, schliefslich aber denselben Wert annimmt, so ist die Anzahl der Vorzeichenwechsel eine gerade. Da nun aber die linke Seite angiebt, wie viel mehr ein Übergang vom Positiven ins Negative, als vom Negativen ins Positive erfolgt, so ist sie gleich 0. Ändert man den Durchlaufungssinn der Kurve in den entgegengesetzten um, so soll die Bahn mit $\overline{S}$ bezeichnet werden, und es gilt dann offenbar die Gleichung

$$\text{(II)} \qquad \mathbf{E}_S\,[A(x, y) : B(x, y)] + \mathbf{E}_{\overline{S}}\,[A(x, y) : B(x, y)] = 0,$$

da mit dem Umlaufungssinn auch die Übergänge ihren Sinn wechseln, d. h. jedem Übergange von $A(x, y) : B(x, y)$ vom Positiven ins Negative längs S ein Übergang vom Negativen ins Positive bei der Umlaufung längs der Bahn $\overline{S}$ und umgekehrt entspricht.

Wenn man auf einer Bahn S zwei Punkte (x_0, y_0) und (x_1, y_1) durch einen beliebigen Weg W mit einander verbindet, so erhält man dadurch zwei geschlossene Bahnen S′ und S″, die den Weg W mit einander gemeinsam haben. Durchläuft man diese beiden Bahnen S′ und S″ so, daſs ihre mit S gemeinsamen Stücke in demselben Sinne wie bei der Bahn S durchlaufen werden, so ist durch diese Festsetzung der Durchlaufungsinn für jede völlig bestimmt, und es wird dann das Stück W bei der Bahn S′ in entgegengesetzter Richtung durchlaufen wie bei S″. Diese Betrachtung kann man verallgemeinern und aus der geschlossenen Bahn S durch fortwährende Einschaltung von Wegen ein System von geschlossenen Bahnen $S_1, S_2, \ldots S_n$ herstellen, deren Durchlaufungsinn völlig bestimmt und so beschaffen ist, daſs jedes nicht mit S gemeinschaftliche Stück bei einer der Bahnen in entgegengesetztem Sinn durchlaufen wird wie bei einer gewissen andern. Bildet man unter dieser Voraussetzung die einzelnen Excesse für die Bahnen $S_1, S_2, \ldots S_n$, so ist klar, daſs sich bei der Summierung diejenigen Teile heben, die sich auf die eingeschalteten Stücke beziehen, und somit nur die übrig bleiben, die sich auf sämtliche Teile der Bahn S erstrecken. Wir können das Resultat so aussprechen:

Zerlegt man eine Bahn S in mehrere Bahnen $S_1, S_2, \ldots S_n$ so ist

$$\text{(III)}\quad E_S[(A(x,y):B(x,y)] = \sum_k E_{S_k}[A(x,y):B(x,y)]. \qquad (k = 1, 2, \ldots n)$$

Eine solche Zerlegung wollen wir nun zur Anwendung bringen, um $E_S[A(x,y):B(x,y)]$ zu bestimmen unter der Voraussetzung, daſs sich die beiden Kurven A und B innerhalb des von der Bahn S eingeschlossenen Gebietes nicht schneiden. Denkt man sich die Kurven A und B gezeichnet, so wird durch sie sowie durch die äuſsere Umgrenzungskurve S das ganze Gebiet in mehrere kleinere Gebiete zerlegt werden. Greifen wir irgend ein Teilgebiet heraus und betrachten seine Umgrenzung, so besteht diese aus Teilen der Kurven A, B und S oder auch nur einigen von ihnen. Niemals können hierbei aber Teile von A und B direkt aneinanderstoſsen, sondern sie müssen immer durch Teile von S von einander getrennt sein. Diesen Umstand benutzen

wir dazu, um das Teilgebiet noch weiter zu zerlegen und zwar auf folgende Weise. Wir fassen irgend einen Teil seiner Begrenzung ins Auge, der zugleich ein Teil von S ist, und verfolgen ihn von einem Punkte α an in einer der beiden möglichen Richtungen, wobei wir schlieſslich zu einem Punkte gelangen, in dem er mit einem Teil von A oder von B zusammenstöſst; wir wollen annehmen mit A. Dann verfolgen wir die Begrenzung weiter bis wir wieder zu einem Teil von S gelangen, denn ein Teil von B kann nicht mit einem Teil von A zusammenstoſsen. So fahren wir weiter fort, bis wir zu einem Teil von S gelangen, der mit einem Teil von B zusammenstöſst. In ersterem nehmen wir dann vorher einen Punkt β an und verbinden ihn durch eine beliebige sich selbst nicht schneidende und ganz innerhalb des Teilgebietes verlaufende Kurve T mit irgend einem Punkte α des Teiles von S, von dem wir ausgegangen waren. Diese Kurve bildet dann mit der durchlaufenen ein geschlossenes Gebiet, auf dessen Begrenzung kein Teil von B liegt, also B(x, y) nicht verschwindet, sondern ein konstantes Vorzeichen hat. Nachdem wir dies Gebiet abgesondert haben, können wir mit dem übrig gebliebenen genau so verfahren wie mit dem Teilgebiet, doch muſs man, wenn die soeben gegebene Beschreibung ohne Modifikation passen soll, die zerlegende Kurve T als einen Teil von S betrachten. Es ist klar, daſs man bei der Fortsetzung dieses Verfahrens und bei der Anwendung auf sämtliche ursprüngliche Teilgebiete schlieſslich dazu gelangt, das ganze von der Kurve S umschlossene Gebiet in solche Teile zu zerlegen, auf deren Begrenzungen entweder A(x, y) oder B(x, y) von konstantem Vorzeichen sind. Deswegen sind aber die über diese Begrenzungen erstreckten Excesse gleich Null, und es folgt somit aus III:

IV. Schlieſst die Bahnkurve S ein Gebiet ein, in dem sich die Kurven A und B nicht schneiden, so ist

$$\mathrm{E_S}\,[A(x, y) : B(x, y)] = 0.$$

Wir wollen ferner nicht verfehlen, ausdrücklich darauf aufmerksam zu machen, daſs sich die Sätze in § 87 mit Leichtigkeit auf Excesse mit geschlossenen Bahnen über-

tragen lassen, und darauf den Beweis des folgenden Satzes gründen, der uns bald von grofsem Nutzen sein wird.

V. Sind A, B, C, D, X, Y Funktionen von x, y, und ist eine der vier ersten, sowie ihre Determinante AD — BC auf der Bahn S von konstantem Vorzeichen, so ist

$$E_S[AX + BY : CX + DY] = \varepsilon E_S(X:Y),$$

wenn ε das konstante Vorzeichen von AD — BC bedeutet.

Für den Beweis legen wir die Annahme zu Grunde, dafs A(x, y) auf S das konstante Vorzeichen ε' hat. Dann ergiebt sich zunächst (§ 87, III)

$$E_S[AX + BY : CX + DY] = \varepsilon' E_S[AX + BY : ACX + ADY].$$

Da nun

$$ACX + ADY = (AD - BC)Y + C(AX + BY)$$

ist, so folgt weiter (§ 87, II)

$$\begin{aligned} &E_S[AX + BY : ACX + ADY] \\ &= E_S[AX + BY : (AD - BC)Y] = \varepsilon E_S[AX + BY : Y] \\ &= -\varepsilon E_S[Y : AX + BY] = -\varepsilon E_S(Y : AX) \\ &= -\varepsilon\varepsilon' E_S(Y : X) = \varepsilon\varepsilon' E_S(X : Y), \end{aligned}$$

wobei aufser § 87, II und III auch noch die Formel I dieses Paragraphen zur Anwendung kommt. Setzt man noch ein, so erhält man sofort das zu beweisende Resultat.

Der Excefs längs einer geschlossenen Bahn hat eine sehr einfache geometrische Bedeutung. Um diese zu entwickeln, betrachten wir zunächst den Exzefs $E_S(x:y)$. Lassen wir den Punkt P mit den Koordinaten x, y die geschlossene Kurve S durchlaufen, so geht er vom ersten Quadranten in den zweiten oder vom dritten in den vierten über, so oft das Verhältnis x : y bei seinem Verschwinden ein Vorzeichenwechsel vom Positiven ins Negative erleidet, während ein Übergang in umgekehrter Richtung erfolgt, also vom zweiten Quadranten in den ersten oder vom vierten in den dritten, so oft x : y beim Verschwinden vom Negativen ins Positive übergeht. Sobald sich die Kurven x = 0 und y = 0 innerhalb des von S umschlossenen Gebietes nicht

schneiden, hat der Excefs den Wert Null nach Satz IV. Schneiden sich aber die beiden Kurven in einem Punkte 0 innerhalb des Gebietes, so kann man um den Punkt 0 einen Kreis K zeichnen und einen beliebigen Punkt α desselben mit einem gleichfalls beliebigen Punkte β von S verbinden durch eine Linie L. Dadurch entsteht ein Gebiet, in dem 0 nicht liegt, und der Excefs über seine Begrenzung verschwindet also. Es bleibt also nur übrig, den Excefs über den Kreis K näher zu betrachten. Dieser ist aber sehr einfach zu ermitteln, denn das Verhältnis $x:y$ geht bei seinem Verschwinden beide Male vom Positiven ins Negative oder vom Negativen ins Positive über, hat also entweder den Wert $+2$ oder -2. Durchläuft man S in entgegengesetzter Richtung, so wird auch der Kreis K in entgegengesetzter Richtung durchlaufen. Nennen wir den Umlauf einen positiven, wenn ein Punkt der Reihe nach die vier Quadranten durchläuft, den entgegengesetzten Umlauf einen negativen, so ist diese Bezeichnung auch auf S übertragbar, und wir können daher kurz sagen, dafs der halbe Excess $E_S(x:y)$ den Umlauf des Punktes (x, y) auf S um den Koordinatenanfangspunkt angiebt.

S haben wir uns bisher immer als eine sich nicht schneidende Kurve vorgestellt. Lassen wir diese Voraussetzung fallen, und nehmen also blofs an, dafs S geschlossen sei, so kann man S in mehrere geschlossene Kurven zerlegen, die sich selbst nicht schneiden. Wendet man die Betrachtung auf jede dieser Kurven an, so gelangt man leicht zu folgendem Satze:

VI. Die Anzahl der Umlaufungen von S um den Koordinatenanfang wird angegeben durch $\frac{1}{2} E_S(x:y)$.

Setzen wir an Stelle von x und y resp. $x-a$ und $y-b$, so ergiebt sich allgemeiner:

Die Anzahl der Umlaufungen von S um den Punkt (a, b) ist gleich $\frac{1}{2} E_S(x-a:y-b)$.

Haben wir nun endlich den Excefs $E_S(X(x, y):Y(x, y))$ vor uns, so können wir X und Y als Koordinaten eines Punktes in einem zweiten Koordinatensystem ansehen.

Durchläuft der Punkt P mit den Koordinaten x, y im ersten Koordinatensystem eine geschlossene Kurve S, so thut es der Punkt mit den Koordinaten X, Y im zweiten System ebenfalls. Durch $\frac{1}{2} E(X:Y)$ wird dann die Umlaufung dieser zweiten Kurve um den Koordinatenanfang ausgedrückt.

---

§ 92. Anzahl der komplexen Wurzeln in einem geschlossenen Gebiet.

Mit Hülfe der entwickelten Sätze sind wir nun imstande, die fundamentale Aufgabe zu lösen, die sich auf die Bestimmung der Anzahl der komplexen Wurzeln einer algebraischen Gleichung in einem geschlossenen Gebiete bezieht. Wir setzen wie früher (§ 83)

$$(1) \quad z = x + iy, \quad F(z) = X(x, y) + iY(x, y).$$

Betrachten wir zunächst den einfachsten Fall, dafs innerhalb des von S umschlossenen Gebietes keine komplexen Wurzeln vorhanden sind, so lehrt uns der Satz IV des vorigen Paragraphen dafs in diesem Falle

$$E_S[X(x, y) : Y(x, y)] = 0$$

ist.

Ist nun z_0 eine Wurzel der Gleichung $F(z) = 0$, so kann man

$$(2) \qquad F(z) = (z - z_0)^m F_1(z)$$

setzen, wo m eine ganze positive Zahl bedeutet, die mindestens den Wert 1 hat, und $F_1(z)$ eine ganze Funktion von z. Man kann m so grofs wählen, dafs $F_1(z)$ nicht die Wurzel z_0 mehr hat, also $F_1(z_0) \neq 0$ wird. Dann nennen wir m den Grad der Wurzel z_0. Da $F_1(z_0) \neq 0$ angenommen ist, so kann man um den Punkt z_0 ein Gebiet so bestimmen, dafs in ihm $F_1(z)$ sich beliebig wenig von $F_1(z_0)$ unterscheidet (§ 84, I), und zwar kann man, wenn

$$(3) \quad F_1(z) = X_0(x, y) + iY_0(x, y), \quad F_1(z_0) = a + ib$$

gesetzt wird, ein Gebiet mit der Umgrenzung S so bestimmen, dafs in diesem sowohl wie auf S $X_0(x, y)$ das

Vorzeichen von a, $Y_0(x, y)$ das Vorzeichen von b hat. Setzen wir nun noch

$$(4) \qquad z^m = (x + iy)^m = X_m(x, y) + i\, Y_m(x, y),$$

also

$$(z - z_0)^m = [(x - x_0) + i(y - y_0)]^m$$
$$= X_m(x - x_0, y - y_0) + i\, Y_m(x - x_0, y - y_0),$$

so kann man auf die Berechnung des Excesses des Verhältnisses der Funktionen

$$(5) \qquad \begin{aligned} X(x, y) &= X_0(x, y)\, X_m(x - x_0, y - y_0) \\ &\quad - Y_0(x, y)\, Y_m(x - x_0, y - y_0) \\ Y(x, y) &= Y_0(x, y)\, X_m(x - x_0, y - y_0) \\ &\quad + X_0(x, y)\, Y_m(x - x_0, y - y_0) \end{aligned}$$

den Satz V des vorigen Paragraphen anwenden, wobei

$$A = D = X_0(x, y), \quad -B = C = Y_0(x, y)$$

ist, also die Determinante

$$A D - B C = X_0^2(x, y) + Y_0^2(x, y)$$

das positive Vorzeichen hat, und man erhält dann

$$(6) \qquad \mathrm{Es}\,[X(x, y) : Y(x, y)]$$
$$= \mathrm{Es}\,[X_m(x - x_0, y - y_0) : Y_m(x - x_0, y - y_0)].$$

Der rechts stehende Exceſs ist nun noch zu berechnen. Wir behandeln zuerst den einfachsten Fall $m = 1$, um darauf zum allgemeinen Falle überzugehen.

I) Ist $m = 1$, und wird die Bahn S in positivem Sinne um den Punkt z_0 umlaufen, so ist nach dem Satze VI des vorigen Paragraphen

$$\frac{1}{2}\,\mathrm{Es}\,(x - x_0 : y - y_0) = 1,$$

und daher giebt auch

$$\frac{1}{2}\,\mathrm{Es}\,[X(x, y) : Y(x, y)]$$

die Anzahl der komplexen Wurzeln innerhalb des Gebietes an, das von der Bahn S umschlossen wird, vorausgesetzt daſs die Umlaufung in positivem Sinne erfolgt.

Wenn nun $F(z)$ zu $F'(z)$ teilerfremd ist, so müssen die Grade aller Wurzeln von $F(z)$ gleich 1 sein. Hat man dann ein beliebiges geschlossenes Gebiet, so kann man es stets so zerlegen, dafs in jedem Teilgebiete entweder eine einzige Wurzel vom ersten Grade oder gar keine vorhanden ist. Ist S die Begrenzung des ganzen Gebietes, so ist der Excefs

$$E_S[X(x, y) : Y(x, y)]$$

gleich der Summe der Excesse über die Begrenzungen der Teilgebiete und daraus folgt sofort, dafs der obige Satz überhaupt für jedes beliebige Gebiet Gültigkeit hat, wenn $F(z)$ und $F'(z)$ teilerfremd sind.

II) Auch im allgemeinen Falle gilt etwas Ähnliches. Hier handelt es sich zunächst darum, den Excefs

$$E_S[X_m(x, y) : Y_m(x, y)]$$

für eine beliebige geschlossene Bahn S zu berechnen, die nicht durch den Koordinatenanfang hindurchgeht, wobei $X_m(x, y)$ und $Y_m(x, y)$ bestimmt sind durch die Gleichung

$$(x + i y)^m = X_m(x, y) + i Y_m(x, y),$$

so dafs allgemein

$$7) \qquad \begin{aligned} X_m &= X_{m-1} x - Y_{m-1} y \\ Y_m &= X_{m-1} y + Y_{m-1} x \end{aligned}$$

ist. Es ist von Wichtigkeit zu bemerken, dafs X_m und Y_m Formen $m$ten Grades von x und y sind, und dafs sich aus Y_m der Faktor y absondern läfst. Wenden wir nun den Satz IV in § 87 an, so ergiebt sich

$$(8) \qquad E_S(X_m : Y_m) = E_S\left(X_m : \frac{Y_m}{y} y\right)$$
$$= E_S\left(X_m y : \frac{Y_m}{y}\right) + E_S\left(\frac{X_m Y_m}{y} : y\right).$$

Da $\frac{X_m Y_m}{y}$ eine Form $(2m - 1)$ten Grades von x, y ist, so können wir sie nach § 87 II durch Weglassung von Potenzen von y auf x^{2m-1} reduzieren, was mit x gleiches Vorzeichen hat; so ergiebt sich, dafs das zweite Glied der rechten Seite gleich $E_S(x : y)$ ist. Um das erste Glied

umzuformen, bemerken wir, daſs zufolge der aus (7) sich ergebenden Gleichung

$$X_m y - Y_m x = -(x^2 + y^2) Y_{m-1}$$

$X_m y$ sich nach dem Modul $\frac{Y_m}{y}$ durch die Funktion

$$-(x^2 + y^2) Y_{m-1}$$

ersetzen läſst, die das entgegengesetzte Vorzeichen Y_{m-1} hat. Daher hat das erste Glied auf der rechten Seite von (8) den Wert $\mathrm{Es}\left(-Y_{m-1} : \frac{Y_m}{y}\right) = -\mathrm{Es}\left(Y_{m-1} : \frac{Y_m}{y}\right)$ $= \mathrm{Es}\left(\frac{Y_m}{y} : Y_{m-1}\right)$. Aus (8) wird also

$$(9) \quad \mathrm{Es}(X_m : Y_m) = \mathrm{Es}(x : y) + \mathrm{Es}\left(\frac{Y_m}{y} : Y_{m-1}\right).$$

Nun wenden wir abermals den Satz IV § 87 an, indem wir Y_{m-1} als Produkt von $\frac{Y_{m-1}}{y}$ und y ansehen, und erhalten so

$$(10) \quad \mathrm{Es}\left(\frac{Y_m}{y} : \frac{Y_{m-1}}{y} y\right) = \mathrm{Es}\left(Y_m : \frac{Y_{m-1}}{y}\right) + \mathrm{Es}\left(\frac{Y_m Y_{m-1}}{y^2} : y\right).$$

Da $\frac{Y_m Y_{m-1}}{y^2}$ sich durch Weglassung von Potenzen von y auf die Potenz x^{2m-3} reduziert, die mit x gleiches Vorzeichen hat, so hat das zweite Glied der rechten Seite wieder den Wert $\mathrm{Es}(x : y)$. Das erste Glied formt man weiter mit Hülfe von § 87 II um. Da nämlich

$$Y_m = Y_{m-1} x + X_{m-1} y$$

ist, so läſst sich Y_m nach dem Modul $\frac{Y_{m-1}}{y}$ auf $X_{m-1} y$ reduzieren, und es ergiebt sich folglich

$$(11) \quad \mathrm{Es}(Y_m : Y_{m-1}) = \mathrm{Es}(x : y) + \mathrm{Es}\left(X_{m-1} y : \frac{Y_{m-1}}{y}\right).$$

Nun haben wir aber oben in (8) schon eine Reduktion

für das zweite Glied der rechten Seite gefunden, nur dafs dort nicht m — 1, sondern m als Index auftritt. Demnach ist

$$(12)\qquad \mathrm{E_S}(X_{m-1}:Y_{m-1}) = \mathrm{E_S}(x:y) + \mathrm{E_S}\left(X_{m-1}y:\frac{Y_{m-1}}{y}\right).$$

Aus den Gleichungen (9) bis (12) folgt nun, dafs

$$(13)\qquad \mathrm{E_S}(X_m:Y_m) = \mathrm{E_S}(x:y) + \mathrm{E_S}(X_{m-1}:Y_{m-1})$$

ist. Setzen wir nun an Stelle von m der Reihe nach m — 1, m — 2, ... 3, 2, so ergiebt sich

$$\mathrm{E_S}(X_{m-1}:Y_{m-1}) = \mathrm{E_S}(x:y) + \mathrm{E_S}(X_{m-2}:Y_{m-2})$$
$$\vdots$$
$$\mathrm{E_S}(X_2:Y_2) = \mathrm{E_S}(x:y) + \mathrm{E_S}(x:y)$$

und daraus durch Addition

$$(14)\qquad \mathrm{E_S}(X_m:Y_m) = m\,\mathrm{E_S}(x:y).$$

In der Formel (6) war nun aber $\mathrm{E_S}[X_m(x-x_0, y-y_0) : Y_m(x-x_0, y-y_0)]$ zu bestimmen. Wir erhalten also, wenn S eine geschlossene Bahn ist, die nicht durch den Punkt (x_0, y_0) hindurchgeht,

$$(15)\qquad \mathrm{E_S}[X_m(x-x_0, y-y_0):Y_m(x-x_0, y-y_0)] = m\,\mathrm{E_S}(x-x_0:y-y_0).$$

Nehmen wir nun an, dafs die Umlaufung von S in positivem Sinne um den Punkt (x_0, y_0) erfolgt, so ist der Excefs auf der rechten Seite gleich 2, und es folgt dann aus (6) unter der dort gemachten Beschränkung über das von S eingeschlossene Gebiet

$$\frac{1}{2}\mathrm{E_S}[X(x,y):Y(x,y)] = \frac{1}{2} m\,\mathrm{E_S}(x-x_0:y-y_0) = m.$$

Es gilt also der in I) gefundene Satz auch dann allgemein, wenn man eine Wurzel vom Grade m als eine m fache Wurzel ansieht.

Wenn wir jetzt das von S umschlossene Gebiet beliebig weit annehmen, unter Beachtung des Satzes II des vorigen Paragraphen, so erhalten wir als Resultat unserer Betrachtungen den folgenden allgemeinen Satz:

Die Anzahl der komplexen Wurzeln einer Gleichung $F(z) = X(x, y) + i\,Y(x, y) = 0$ in einem beliebigen von einer geschlossenen Bahn S umgrenzten Gebiete wird durch den halben um das Gebiet in positivem Umlauf erstreckten Exceſs

$$\frac{1}{2}\,\mathrm{E}_S\,[X(x, y) : Y(x, y)]$$

ausgedrückt. Dieser Satz ist zuerst von Cauchy aufgestellt und bewiesen worden. Mit Rücksicht auf die Bemerkung am Schlusse des vorhergehenden Paragraphen kann man dem Satze auch die folgende einfache Fassung geben:

Die Anzahl der komplexen Wurzeln der Gleichung $F(z) = X(x, y) + i\,Y(x, y) = 0$ in einem begrenzten Gebiet ist gleich der Anzahl der Umläufe, des Punktes (X, Y) um den Nullpunkt, die man erhält, wenn man den Punkt (x, y) in positivem Umlauf um das Gebiet herumführt.

---

§ 93. Fundamentalsatz der Algebra.

Wir wollen jetzt die entwickelte Theorie anwenden, um die Gesamtzahl der reellen oder complexen Wurzeln der algebraischen Gleichung

$$F(z) = c_0 z^n + c_1 z^{n-1} + \ldots + c_n = 0$$

vom Grade n zu ermitteln.

Wie früher (§ 84) bewiesen ist, kann man stets eine Gröſse r so bestimmen, daſs für alle Werte von z, deren absoluter Betrag gröſser als r ist, $|F(z)| > \omega$ wird, wo ω eine beliebige positive Gröſse darstellt, die man so groſs annehmen kann, als man will. Alle die genannten Werte von z lassen sich geometrisch sehr einfach charakterisieren: die ihnen entsprechenden Punkte liegen nämlich alle auſserhalb des Kreises mit den Radius r um den Koordinatenanfang als Mittelpunkt. Auſserhalb dieses Kreises kann es keine Wurzeln von $F(z) = 0$ geben, und es kommt somit die Frage

nach der Anzahl aller überhaupt vorhandenen Wurzeln auf die speziellere zurück, wie viele Wurzeln innerhalb des genannten Kreises liegen. Wir entscheiden dies, indem wir den Exzefs des der Funktion

$$F(z) = X(x, y) + i\,Y(x, y)$$

entsprechenden Funktionenverhältnisses $X(x, y) : Y(x, y)$ längs eines positiven Umlaufes um die Kreisbahn S berechnen.

Wenn

$$F(z) = z^n F_0(z)$$

$$F_0(z) = c_0 + \frac{c_1}{z} + \dots + \frac{c_n}{z^n} = X_0(x, y) + i\,Y_0(x, y),$$

und wie im vorigen Paragraphen

$$z^n = (x + i\,y)^n = X_n(x, y) + i\,Y_n(x, y)$$

gesetzt wird, so ergiebt sich aus

$$X = X_0 X_n - Y_0 Y_n, \quad Y = Y_0 X_n + X_0 Y_n$$

mit Hilfe des Satzes V § 91, da $X_0^2 + Y_0^2 > 0$ ist, X_0 und Y_0 constantes Vorzeichen haben, und der Formel (15) des vorigen Paragraphen

$$\mathbf{E}_S[X(x, y) : Y(x, y)] = \mathbf{E}_S[X_n(x, y) : Y_n(x, y)] = n\,\mathbf{E}_S(x : y).$$

Wenn aber der Umlauf einmal und in positivem Sinne erfolgt, so ist $\mathbf{E}_S(x : y) = 2$, also $\mathbf{E}_S(X : Y) = 2n$. Wir erhalten somit den Satz:

Jede algebraische Gleichung hat genau soviel reelle oder komplexe Wurzeln, wie ihr Grad angiebt.

Dieser Satz wird häufig als der Fundamentalsatz der Algebra bezeichnet. Er ist in der That von ungeheurer Wichtigkeit, da er zeigt, dafs mit Einführung des Begriffes komplexer Wurzeln ein gewisser Abschlufs der Algebra herbeigeführt wird, da es dann keine Gleichung mehr giebt, die unlösbar ist.

Hat die Gleichung $F(z) = 0$ eine Wurzel z_1 vom Grade m_1, z_2 vom Grade m_2, $\dots z_k$ vom Grade m_k und keine weiteren, so ist

$$n = m_1 + m_2 + \dots m_k$$

$$F(z) = c_0 (z - z_1)^{m_1} (z - z_2)^{m_2} \dots (z - z_k)^{m_k}.$$

Es läfst sich also jede ganze rationale Funktion als ein Produkt von lauter Linearfaktoren darstellen.

Alles dieses gilt, ob nun die Koeffizienten von $F(z)$ reell oder komplex sind. Sind sie aber reell, so kann man den Fundamentalsatz der Algebra eine von dem Begriffe der komplexen Gröfsen unabhängige Fassung geben. Wird nämlich

$$F(x + i\,y) = X(x, y) + i\,Y(x, y),$$

so ist in diesem Falle auch

$$F(x - i\,y) = X(x, y) - i\,Y(x, y),$$

wovon man sich leicht überzeugt. Hat also die Gleichung $F(z)$ die Wurzel $x_0 + i\,y_0$, so hat sie auch die dazu konjugiert komplexe Gröfse $x_0 - i\,y_0$ als Wurzel. Betrachtet man nun das Produkt

$$[z - (x_0 + i\,y_0)]\,[z - (x_0 - i\,y_0)] = (z - x_0)^2 + y_0^2$$

zweier ihnen entsprechenden Linearfaktoren, so ergiebt sich eine quadratische Funktion

$$z^2 - 2\,x_0\,z + x_0^2 + y_0^2.$$

Mit Rücksicht auf diesen Umstand kann man daher unter alleiniger Beibehaltung des Begriffes irrationaler Gröfsen den Fundamentalsatz der Algebra in der Form ausdrücken:

Jede ganze rationale Funktion mit reellen Koeffizienten läfst sich als ein Produkt von lauter Funktionen ersten oder zweiten Grades mit reellen Koeffizienten darstellen.

XI. Abschnitt.

Näherungsmethoden.

§ 94. Begriff und Zweck der Näherungsmethoden.

Die im vorhergehenden Abschnitte dargelegten Methoden reichen zwar völlig aus, um die reellen Wurzeln der Gleichungen in beliebig kleine Intervalle einzuschliefsen und daher so genau, als man will, zu berechnen. Sie sind ferner auch wirklich zweckmäfsig, so lange es sich noch um die Bestimmung grofser Intervalle handelt. Hat man aber schon ein kleines Intervall ermittelt und will dieses weiter verengen, so kommt man doch nur verhältnismäfsig langsam zum Ziele, denn, halbiert man fortgesetzt das Intervall, was am schnellsten vorwärts führt, so ist die Gröfse des ursprünglichen Intervalles von der Breite d nach n Schritten erst auf $\frac{d}{2^n}$ herabgesunken. Man hat daher dann früh zu Methoden seine Zuflucht zu nehmen, die weitaus schneller zum Ziele führen. Von diesen Näherungsmethoden wollen wir die wichtigsten jetzt darlegen.

§ 95. Newtonsche Näherungsmethode.

Die Newtonsche Näherungsmethode setzt voraus, dafs man ein Intervall a...b von der Gröfse d bestimmt hat, in dem F(x) sein Vorzeichen nur einmal wechselt, die Derivierten F′(x) und F″(x) die ihrigen dagegen beibehalten. Diese Voraus-

setzung läfst sich immer realisieren, wenn $F(x)$ und $F'(x)$ teilerfremd sind; wenn aber dies nicht der Fall ist, so läfst sich $F(x)$, wie wir früher gesehen haben (§ 65), immer in Faktoren zerfällen, deren Derivierten zu ihnen teilerfremd sind, und diese hat man dann für sich einzeln zu untersuchen.

Die Newtonsche Methode lehrt nun ein in dem Intervall $a \ldots b$ enthaltenes kleineres Intervall $a' \ldots b'$ bestimmen, für das die sämtlichen genannten Voraussetzungen auch noch zutreffen, und zwar auf folgende Weise: Nennen wir a und a' die Endpunkte der Intervalle, in denen $F(x)$ dasselbe Vorzeichen besitzt wie $F''(x)$, b und b' die andern Endpunkte, in denen also das Vorzeichen von $F(x)$ das entgegengesetzte ist, so ist

$$a' = a - \frac{F(a)}{F'(a)}, \quad b' = b - \frac{F(b)}{F'(a)}.$$

Bevor wir den Beweis dafür erbringen, dafs a' und b' im Innern des Intervalles liegen und die Wurzel der Gleichung $F(x) = 0$ einschliefsen, wollen wir die einfache Überlegung angeben, die der Methode historisch zu Grunde liegt. Ist nämlich a ein angenäherter Wert der Wurzel x_0 und entwickelt man $F(x_0)$ nach Potenzen von $(x_0 - a)$, beschränkt sich dabei auf die erste Potenz, so ergiebt sich aus der angenäherten Gleichung

$$F(a) + (x_0 - a) F'(a) + \ldots = F(x_0) = 0$$

der Näherungswert

$$x_0 = a - \frac{F(a)}{F'(a)},$$

der oben mit a' bezeichnet wurde.

Soll diese Schlufsweise exakt gestaltet werden, so kann man den Mittelwertsatz der Differentialrechnung anwenden. (S. Bd. X dieser Sammlung.)

Man kann dann setzen

$$F(x) = F(a) + (x - a) F'(\bar{a})$$
$$F(x) = F(b) + (x - b) F'(\bar{b}),$$

wobei $\bar{a}$ und $\bar{b}$ noch unbekannte Werte sind, von denen aber soviel feststeht, dafs der erstere im Intervalle $x \ldots a$,

der zweite im Intervalle $x \ldots b$ liegt. Setzen wir für x die Wurzel x_0, so folgt

$$x_0 = a - \frac{F(a)}{F'(\bar{a}_0)}$$

$$x_0 = b - \frac{F(b)}{F'(\bar{b}_0)},$$

wenn $\bar{a}_0$ im Intervalle $x_0 \ldots a$, b_0 im Intervalle $x_0 \ldots b$ liegt. Von den beiden Werten für x_0 gebrauchen wir den ersten zur Berechnung von $\frac{a' - a}{x_0 - a'}$, den zweiten setzen wir in $\frac{b' - b}{x_0 - b'}$ ein und erhalten so

$$\frac{a' - a}{x_0 - a'} = \frac{F'(\bar{a}_0)}{F'(a) - F'(\bar{a}_0)}$$

$$\frac{b' - b}{x_0 - b'} = \frac{F'(\bar{b}_0)}{F'(a) - F'(\bar{b}_0)}.$$

Die hinreichende und notwendige Bedingung dafür, dafs a' im Intervall $x_0 \ldots a$, b' im Intervall $x_0 \ldots b$ liegt, besteht nun darin, dafs die vorliegenden Ausdrücke positiv sind. Dafs dem wirklich so ist, ergiebt sich aus den Annahmen. $F''(x)$ wechselt im Intervall $a \ldots b$ das Vorzeichen nicht; nach einem bekannten Satz der Differentialrechnung mufs daher $F'(x)$ sich im Intervall beständig in demselben Sinne ändern, d. h. entweder ab- oder zunehmen, wenn x das Intervall durchläuft, so dafs für jeden Wert von x im Intervall $a \ldots b$ das Vorzeichen der drei Gröfsen

$$F'(a) - F'(x),\ F'(x) - F'(b),\ (a - b)\,F''(x)$$

dasselbe ist. Ferner ist nach den Annahmen das Vorzeichen von $F(a)$ übereinstimmend mit dem von $F''(x)$. Bedenken wir nun noch, dafs stets das Vorzeichen von $F(a)$ auch mit dem von $(a - b)\,F'(x)$ identisch ist, so erkennen wir, dafs die drei Gröfsen

$$F'(a) - F'(x),\ F'(x) - F'(b),\ F'(x)$$

dasselbe Vorzeichen haben, oder dafs

$$\frac{F'(x)}{F'(a) - F'(x)} > 0, \quad \frac{F'(x)}{F'(x) - F'(b)} > 0$$

ist, wenn x ein Intervall a ... b liegt, und wir brauchen nun bloſs noch für dieses x die Werte $\overline{a_0}$, $\overline{b_0}$ zu setzen, um die Richtigkeit unserer Behauptung zu erkennen.

Um die Gröſse d' des neuen Intervalles zu bestimmen und so eine Einsicht in die Art der Annäherung zu gewinnen, berechnen wir

$$b' - a' = b - a - \frac{F(b) - F(a)}{F'(a)}$$

unter Anwendung des Mittelwertsatzes in der Form

$$F(b) = F(a) + (b - a) F'(a) + \frac{1}{2}(b - a)^2 F''(c),$$

wo wieder c einen im Innern des Intervalles a ... b gelegenen Wert bedeutet. Es ergiebt sich so

$$\pm d' = b' - a' = -\frac{1}{2}(b - a)^2 \frac{F''(c)}{F'(a)}.$$

Bezeichnen wir nun mit M den absoluten Maximalwert von $F''(x)$, mit m den absoluten Minimalwert von $F'(x)$ im Intervalle a ... b und setzen noch

$$\frac{1}{2}\frac{M}{m} = e,$$

so zeigt sich, daſs

$$d' \leqq e\, d^2.$$

Wenn man die Näherungsmethode nun n mal angewandt hat, so findet sich für die Gröſse der weiteren Intervalle $d'', d''', \ldots d^{(n)}$

$$d'' < e' d'^2, \quad d''' < e'' d''^2, \ldots d^{(n)} < e^{(n-1)} d^{(n-1)^2},$$

und hierbei ist

$$e' \leqq e, \quad e'' \leqq e', \ldots e^{(n-1)} \leqq e^{(n-2)} \leqq e,$$

so daſs

$$d^{(n)} \leqq e^{2^n - 1} d^{2^n},$$

woraus hervorgeht, daſs die Intervalle auſserordentlich schnell abnehmen, wenn $d < 1$ ist.

Da man die erreichte Genauigkeit auf diese Weise leicht ermessen kann, so ist es überflüssig, die Reihe der Werte b', b'', ... noch besonders zu berechnen, was mit großem Arbeitsaufwand verbunden wäre.

Ein Beispiel möge die Anwendung der Methode veranschaulichen. Die Gleichung

$$F(x) = x^3 + x^2 - 4x + 1 = 0$$

hat, wie man leicht erkennt, eine Wurzel zwischen 0,2 und 0,3, und innerhalb dieser Grenzen sind auch die Voraussetzungen für die Anwendbarkeit der Newtonschen Methode erfüllt; die Gröſse e ist kleiner als 1, die Näherungswerte sind also auf die 2te, 4te, 8te... Dezimalstelle genau. Man erhält folgende Tafel

a	$F(a)$	$F'(a)$	$-\frac{F(a)}{F'(a)}$
0,2	0,248	— 3,48	0,07
0,27	0,012 583	— 3,2413	0,0038
0,2738	0,000 292 251 272	— 3,227 500 68	0,000 090 55

und damit den auf 8 Dezimalstellen sichern Wurzelwert 0,273 890 55.

§ 96. Lagrangesche Näherungsmethode.

Die Lagrangesche Näherungsmethode führt die Bestimmung der Wurzeln auf die Bildung einer Reihe von transformierten Gleichungen zurück, von denen die ganzzahligen Näherungswerte der Wurzeln in Betracht gezogen werden. Die Wurzeln der ursprünglichen Gleichung lassen sich dabei als Kettenbrüche mit lauter positiven Teilnennern darstellen (§ 13).

Hat die Gleichung $F(x) = 0$ (eine oder mehrere) Wurzeln zwischen den beiden ganzen Zahlen a und $a + 1$, und setzt man

$$x = a + \frac{1}{x_1}$$

$$F_1(x_1) = x_1^n F\left(a + \frac{1}{x_1}\right) = F(a)\,x_1^n + F'(a)\,x_1^{n-1} + \ldots + \frac{F^{(n)}(a)}{n!},$$

so hat die transformierte Gleichung $F_1(x_1) = 0$ (eine oder mehrere) Wurzeln, die positiv und gröſser als 1 sind. Diese

Gleichung behandelt man wie die vorige, man bestimmt also einen ganzzahligen positiven Wert a_1, zwischen dem und $a_1 + 1$ eine Wurzel von $F_1(x_1)$ liegt u. s. w. Fährt man so fort, indem man immer

$$x_i = a_i + \frac{1}{x_{i+1}} \qquad (i = 1, 2, \ldots)$$

$$F_{i+1}(x_{i+1}) = x_{i+1}^n F_i\left(a_i + \frac{1}{x_{i+1}}\right)$$

setzt, wobei angenommen wird, dafs die Gleichung $F_i(x_i)$ eine in dem Intervalle $a_i \ldots a_i + 1$ gelegene Wurzel hat, so ergiebt sich bei n maliger Anwendung des Verfahrens, dafs

$$x = a + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cdots \cfrac{1}{a_{n-1} + \cfrac{1}{x_n}}}}$$

ist, wo x_n eine Wurzel der Gleichung $F_n(x_n) = 0$ bezeichnet, die positiv und gröfser als 1 ist. Aus der Theorie der Kettenbrüche (§ 13) ergiebt sich dann, wenn man mit P_i und Q_i die Zähler und Nenner des i ten Näherungsbruches bezeichnet, dafs sich der wahre Wert der Wurzel von $\frac{P_n}{Q_n}$ um weniger als $\frac{1}{Q_n^2}$ unterscheidet, und diese Gröfse kann beliebig klein gemacht werden, wenn man das beschriebene Verfahren weit genug fortsetzt, da alle Teilnenner $a_1, a_2, \ldots$ positiv sind, Q_n also über alle Grenzen hinauswächst. Insbesondere läfst sich das Verfahren auch so weit fortsetzen, dafs eine Wurzel völlig isoliert wird, d. h. dafs $F(x_n)$ nur eine einzige positive Wurzel besitzt.

Bei der Anwendung dieses Verfahrens erscheint zunächst die Ermittelung der beiden ganzen Zahlen, zwischen denen die Wurzeln der transformierten Gleichung $F_k(x_k) = 0$ gelegen sind, und die man durch die Vorzeichenwechsel von $F_k(x_k)$ feststellen kann, als der umständlichste Teil der Untersuchung. Wie wir nun zeigen wollen, ist dies nur zuerst der Fall, während nachher eine einfache Regel die Zahlen des Intervalles ermitteln lehrt. Es kommt somit als haupt-

sächlichste Unbequemlichkeit nur die Bildung der transformierten Gleichungen ernstlich in Betracht, da die Berechnung der Näherungswerte keinen weiteren Schwierigkeiten unterliegt.

Die Variabeln x in F(x) und x_k in $F_k(x_k)$ hängen durch die Substitution

$$x = \frac{P_k x_k + P_{k-1}}{Q_k x_k + Q_{k-1}}$$

zusammen, aus der umgekehrt

$$x_k = -\frac{P_{k-1} - Q_{k-1} x}{Q_k x - P_k} = \frac{Q_{k-1}}{Q_k} \frac{\frac{P_{k-1}}{Q_{k-1}} - x}{x - \frac{P_k}{Q_k}}$$

folgt. Jeder Wurzel der Gleichung $F(x) = 0$ entspricht eine solche der Gleichung $F_k(x_k) = 0$ und umgekehrt. Nennen wir die sämtlichen Wurzeln von $F(x) = 0$ $\alpha_1, \alpha_2, \ldots \alpha_n$ und die entsprechenden von $F_k(x_k) = 0$ $\alpha_1^{(k)}, \alpha_2^{(k)}, \ldots \alpha_n^{(k)}$, und nehmen wir an, dafs F(x) und $F_k(x_k)$ als Koeffizienten der höchsten Potenzen die Gröfsen c_0 und $c_0^{(k)}$ haben, so ist

$$F(x) = c_0 \prod_h (x - \alpha_h)$$
$$F_k(x_k) = c_0^{(k)} \prod_h (x_k - \alpha_h^{(k)}). \qquad (h = 1, 2, \ldots n)$$

Es ist nun nicht schwer eine Form zu bestimmen, der sich $F_k(x_k)$ mit wachsendem k nähert. Ist α eine Wurzel der Gleichung $F(x) = 0$, deren Wert durch die Näherungsbrüche gegeben wird, so sind die beiden Differenzen

$$\frac{P_{k-1}}{Q_{k-1}} - \alpha, \quad \frac{P_k}{Q_k} - \alpha$$

von entgegengesetztem Vorzeichen, und es wird daher die entsprechende Wurzel von $F_k(x_k) = 0$

$$\alpha^{(k)} = \frac{Q_{k-1}}{Q_k} \frac{\frac{P_{k-1}}{Q_{k-1}} - \alpha}{\alpha - \frac{P_k}{Q_k}}$$

bei hinreichend grofsen Werten von k positiv sein. Alle übrigen Wurzeln α verhalten sich dagegen völlig anders,

denn die entsprechenden Wurzeln $\alpha^{(k)}$ unterscheiden sich beliebig wenig von der negativen Gröfse $-\frac{Q_{k-1}}{Q_k}$. Dies gilt auch dann noch, wenn $\alpha = p + qi$ komplex wird, denn dann wird

$$\alpha^{(k)} = \frac{Q_{k-1}}{Q_k} \cdot \frac{\frac{P_{k-1}}{Q_{k-1}} - p - qi}{p + qi - \frac{P_k}{Q_k}}$$

$$= -\frac{Q_{k-1}}{Q_k} - \frac{Q_{k-1}}{Q_k} \frac{\left(\frac{P_k}{Q_k} - \frac{P_{k-1}}{Q_{k-1}}\right)\left(p - iq - \frac{P_k}{Q_k}\right)}{\left(p - \frac{P_k}{Q_k}\right)^2 + q^2}.$$

Hat nun $F_k(x_k)$ lauter Wurzeln dieser letzteren Art, so wird $F_k(x_k)$ sich wenig unterscheiden von

$$c_0^{(k)}\left(x_k + \frac{Q_{k-1}}{Q_k}\right)^n.$$

Ist dagegen eine Wurzel α immer zwischen je zwei aufeinander folgenden Brüchen der Reihe

$$\frac{P_1}{Q_1}, \frac{P_2}{Q_2}, \frac{P_3}{Q_3}, \ldots$$

gelegen, so wird $F_k(x_k)$ sich der Form

$$c_0 (x_k - \alpha^{(k)}) \left(x_k + \frac{Q_{k-1}}{Q_k}\right)^{n-1}$$

nähern. Der Koeffizient $-c_1^{(k)}$ der $(n-1)$ Potenz von x_k in $F_k(x_k)$ wird also nahezu gleich $c_0^{(k)}\left[(n-1)\frac{Q_{k-1}}{Q_k} - \alpha^{(k)}\right]$ sein, oder es wird $\alpha^{(k)}$ sich nicht sehr von $(n-1)\frac{Q_{k-1}}{Q_k} + \frac{c_1^{(k)}}{c_0^{(k)}}$ unterscheiden, und die gröfste hierin enthaltene ganze Zahl als Teilnenner a_k für die folgende Annäherung benutzt werden können, wenn k einen hinreichend grofsen Wert hat.

Beispiel. Man erkennt leicht, dafs die Gleichung

$$F(x) = x^3 + x^2 - 4x + 1 = 0$$

drei reelle Wurzeln und zwar zwischen -3 und -2, 0 und 1, 1 und 2 hat, weil

$$F(-3) = -5, \quad F(-2) = 5; \quad F(0) = 1, \quad F(1) = -1, \quad F(2) = 5$$

ist. Für die Berechnung der zwischen 0 und 1 gelegenen Wurzel ergeben sich folgende Gleichungen, Teilnenner und Näherungsbrüche:

k		$\left[\frac{2Q_{k-1}}{Q_k} + \frac{c_1^{(k)}}{c_0^{(k)}}\right]$	a_k	$P_{k+1} : Q_{k+1}$
0	$x^3 + x^2 - 4x + 1$		0	0:1
1	$x^3 - 4x^2 + x + 1$		3	1:3
2	$5x^3 - 4x^2 - 5x - 1$	1	1	1:4
3	$5x^3 - 2x^2 - 11x - 5$	1	1	2:7
4	$13x^3 - 13x - 5$	2	1	3:11
5	$5x^3 - 26x^2 - 39x - 13$	6	6	20:73
6	$103x^3 - 189x^2 - 64x - 5$	2	2	43:157
7	$65x^3 - 416x^2 - 429x - 103$	7	7	321:1172
8	$1195x^3 - 3302x^2 - 949x - 65$	3	3	1006:3673
9	$365x^3 - 9504x^2 - 7453x - 1195$	26	26	26477:96670

§ 97. Bernoullische und Gräffesche Näherungsmethode.

Die beiden Näherungsmethoden, die wir nun noch besprechen wollen, haben das Gemeinsame, dafs sie unmittelbar nur die Werte der absolut gröfsten und absolut kleinsten Wurzeln einer Gleichung bestimmen lehren mit Hülfe von symmetrischen Funktionen aller Wurzeln der Gleichung, die sich aber rational durch die Koeffizienten darstellen lassen. Das ist in Wirklichkeit aber keine Beschränkung, wie wir zunächst zeigen wollen.

Hat man nämlich einen Wert a ermittelt, der der

Wurzel x_0 der Gleichung $F(x)=0$ näher liegt als jeder der andern Wurzeln x_i, so daſs also

$$|x_0 - a| < |x_i - a| \qquad (i = 1, 2, \dots n-1)$$

ist, so führt die Substitution

$$x' = \frac{1}{x-a}, \qquad x = a + \frac{1}{x'}$$

$$x'^n F\left(a + \frac{1}{x'}\right) = G(x')$$

zu einer Gleichung $G(x')=0$, deren absolut gröſste Wurzel

$$x_0' = \frac{1}{x_0 - a}$$

ist. Da dann

$$x_0 = a + \frac{1}{x_0'}$$

ist, so läſst sich also jede Wurzel x_0 der ursprünglichen Gleichung durch ein einfaches Transformationsverfahren finden.

Sind $x_0, x_1, \dots x_{n-1}$ die sämtlichen Wurzeln der Gleichung $F(x)=0$, und setzen wir

$$s_k = x_0^k + x_1^k + \dots x_{n-1}^k$$

so ist, wenn x_0 die absolut gröſste Wurzel ist,

$$x_0 = \lim_{k=\infty} \frac{s_{k+1}}{s_k}, \qquad x_0 = \lim_{k=\infty} \sqrt[k]{s_k}.$$

Man kann also x_0 angenähert berechnen, wenn man für k einen sehr groſsen Wert nimmt, vorausgesetzt daſs die Gröſsen s_k bekannt sind. Das folgt zwar aus der im folgenden Abschnitt betrachteten Theorie der symmetrischen Funktionen allgemein. Wir brauchen diese hier aber nicht vorauszusetzen, da hier spezielle Methoden zur Anwendung gelangen, die verschieden sind, je nachdem die erste Formel, die die Grundlage der Bernoullischen, oder die zweite, die die der Gräffeschen Näherungsmethode bildet, benutzt wird.

Wie grofs der Genauigkeitsgrad ist, beurteilt man am besten nachträglich durch Einsetzen und Variieren des Näherungswertes, da eine allgemeine Bestimmung nicht zu einfachen Kriterien führt.

I. Bernoullische Methode. Ist

$$F(x) = x^n - a_1 x^{n-1} + a_2 x^{n-2} - \dots \pm a_n,$$

so kann man sich zur Berechnung von s_k der Rekursionsformeln bedienen

$$s_k = a_1 s_{k-1} - a_2 s_{k-2} + \dots \mp a_n s_{k-n}, \qquad (k \geqq n)$$

wenn zuvor die n Gröfsen

$$s_0 = n,\ s_1,\ s_2, \dots s_{n-1}$$

bestimmt sind, was an und für sich nicht mit Schwierigkeiten verknüpft sein würde (§ 100), hier aber völlig umgangen werden kann. Ist nämlich

$$\varphi(x) = b_0 x^{n-1} + b_1 x^{n-2} + \dots + b_{n-1}$$

eine Funktion (n — 1) ten Grades, so kann man den Quotienten

$$\frac{\varphi(x)}{F(x)} = \sum_i \frac{\varphi(x_i)}{(x - x_i) F'(x_i)}, \quad (i = 0, 1, \dots n-1)$$

wenn man

$$\frac{1}{x - x_i} = \frac{1}{x\left(1 - \frac{x_i}{x}\right)} = \sum_k \frac{x_i^k}{x^{k+1}} \qquad (k = 0, 1, \dots \infty)$$

einsetzt, in der Form entwickeln

$$\sum_k \frac{s_k'}{x^{k+1}}, \qquad (k = 0, 1, \dots \infty)$$

wobei

$$s_k' = \sum_i \frac{\varphi(x_i)\, x_i^k}{F'(x_i)} \qquad (i = 0, 1, \dots n-1)$$

ist. Multipliziert man nun beide Seiten der Gleichung

$$\frac{\varphi(x)}{F(x)} = \sum_k \frac{s_k'}{x^{k+1}} \qquad (k = 0, 1, \dots \infty)$$

mit F(x), so erhält man durch Vergleichung der Koeffizienten gleich hoher Potenzen von x auf beiden Seiten folgende Reihe von Gleichungen

$$s_0' = b_0, \; s_1' - a_1 s_2' = b_1, \; s_2' - a_1 s_1' + a_2 s_0' = b_2, \ldots$$

und ferner für $k \geqq n$

$$s_k' - a_1 s_{k-1}' + a_2 s_{k-2}' - \ldots a_n s_{k-n}' = 0.$$

Aus den ersten n Gleichungen ergiebt sich, daſs man zu ganz willkürlichen Werten von $s_0', s_1', \ldots s'_{n-1}$ stets eine zugehörige Funktion $\varphi(x)$ bestimmen kann, die letzteren lehren dann, daſs die Gröſsen s_k' nach einer genau mit der Berechnung von s_k übereinstimmenden Rekursionsformel ermittelt werden können. Überdies ist auch noch

$$x_0 = \lim_{k=\infty} \frac{s_{k+1}'}{s_k'}$$

unter der Voraussetzung, daſs $\varphi(x_0) \neq 0$ ist, die durchaus zulässig ist, da sonst F(x) und $\varphi(x)$ einen Faktor gemeinsam hätten.

Die Willkürlichkeit der Gröſsen $s_0', s_1', \ldots s'_{n-1}$ hat nun aber noch einen zweiten Vorteil, wenn man schon einen angenäherten Wert a der gröſsten Wurzel x_0 kennt. Man kann dann nämlich

$$s_0' = 0, \; s_1' = 0, \ldots s'_{n-2} = 1, \; s'_{n-1} = a$$

setzen und dadurch eine schnellere Annäherung in der Rechnung erzielen.

Als Beispiel betrachten wir die Gleichung

$$x^3 + x^2 - 4x + 1 = 0,$$

deren absolut gröſster Wurzelwert zwischen -2 und -3 liegt. Wir gehen aus von den Werten

$$s_0' = 0, \; s_1' = 1, \; s_2' = -2$$

und berechnen dann mit Hülfe der Gleichung

$$-s'_{k-3} + 4 s'_{k-2} - s'_{k-1} = s_k'$$

die Gröſsen $s_3', s_4', \ldots$, was bei Anwendung geeigneter Rechenmethoden in diesem einfachen Fall sogar ohne

jede Nebenrechnung geschehen kann. Man erhält so die Zahlenreihe

$$6, -15, 41, -107, 286, -755, 2006, -5312, 14091,$$
$$-37345, 99021, -262492, 695921,$$

und es ergiebt sich aus den beiden letzten Zahlen durch Division die Zahl $-2,651\ldots$ als ein Wurzelwert, der bis auf die dritte Dezimalstelle genau ist.

II) Gräffesche Methode. Bei Anwendung der Formel

$$x_0 = \lim_{k=\infty} \sqrt[k]{s_k}$$

berechnet man, um möglichst schnell und einfach zu möglichst hohen Werten von k zu gelangen, die Gröſsen s_k der Reihe nach für die Werte $k = 1, 2, 2^2, 2^3 \ldots$, und zwar in folgender Weise. Da

$$s_2 = x_0^2 + x_1^2 + \ldots + x_n^2$$

ist, so läſst sich s_2 als Wurzelsumme einer transformierten Gleichung $F_1(x') = 0$ mit den Wurzeln $x_0^2, x_1^2, \ldots x_{n-1}^2$ auffassen, die sich durch Elimination von x aus den beiden Gleichungen

$$F(x) = 0, \quad x^2 = x'$$

so bilden läſst: Wir zerlegen $F(x)$ in zwei Summanden, von denen der eine $\varphi(x^2)$ alle geraden, der andere $\psi(x^2)x$ alle ungeraden Potenzen von x enthält, so daſs also

$$F(x) = \varphi(x^2) + \psi(x^2)\,x$$

ist. Dann erhält man sofort als transformierte Funktion

$$F_1(x') = \psi(x')^2\,x' - \varphi(x')^2.$$

So kann man weiter fortfahren. Heiſst die Gleichung, deren Wurzeln $x_0^{2^k}, x_1^{2^k}, \ldots x_{n-1}^{2^k}$ sind,

$$F_k(x^{(k)}) = x^{(k)^n} - a_1^{(k)}\,x^{(k)^{n-1}} + \ldots,$$

so ist

$$x_0 = \lim_{k=\infty} \sqrt[2^k]{a_1^{(k)}}$$

durch fortgesetztes Quadratwurzelausziehen zu erhalten.

Wir betrachten wieder das frühere Beispiel

$$x^3 + x^2 - 4x + 1 = 0$$

und erhalten der Reihe nach folgende transformierte Gleichungen

$$x^3 - 9x^2 + 14x - 1 = 0, \quad x^3 - 53x^2 + 178x - 1 = 0$$

$$x^3 - 2453x^2 + 31578x - 1 = 0,$$

$$x^3 - 5954053x^2 + 997\,165\,178x - 1 = 0.$$

Aus der letzteren ergiebt sich

$$\sqrt[32]{5954053} = 2,651\ldots,$$

und da die Wurzel negativ ist, $-2,651\ldots$ als angenäherter Wurzelwert.

XII. Abschnitt.

Algebraische Auflösung der Gleichungen.

---

§ 98. Das Problem der algebraischen Auflösung und seine historische Entwickelung.

Sobald die Koeffizienten einer Gleichung als reelle oder komplexe Zahlen gegeben sind, ist es, wie wir im vorigen Abschnitte sahen, stets möglich, ihre reellen und komplexen Wurzeln so genau, als man will, zu berechnen. Unterwirft man die Zahlenwerte irgend welchen Änderungen, so müssen jedoch die dargelegten Methoden wieder von Grund auf an mit den neuen Zahlenwerten zur Anwendung gebracht werden. Die Methoden haben also keinen Wert, wenn man die Koeffizienten als unbestimmte Gröfsen ansieht. Daher ist es von Wichtigkeit, zu untersuchen, ob es nicht Auflösungsmethoden giebt, die auf Gleichungen mit unbestimmten Koeffizienten Anwendung finden, so dafs erst zu allerletzt die Einsetzung der Zahlenwerte der Koeffizienten bei der Ermittelung der Wurzeln der Gleichung nötig wird. Solche Methoden giebt es nun in der That.

Schon im Altertum war es bekannt, dafs die Lösung der quadratischen Gleichungen auf die Ausziehung einer Quadratwurzel aus einem rationalen Ausdruck zurückgeführt werden kann; Diophant hat zuerst die Lösung gefunden, die sich aus einigen Sätzen von Euklid naturgemäfs ergiebt. Nach dem Wiederaufleben der Wissenschaften beim Beginne der Neuzeit fanden dann italienische Mathematiker, Scipio Ferreo und Ferrari, dafs auch die Auflösung der kubischen und biquadratischen Gleichungen auf eine

Reihe von Wurzelausziehungen reduziert werden kann, und zwar auf Quadrat- und Kubikwurzeln. Zu diesen Auflösungen gelangte man ursprünglich durch reine Versuche, die, so geringe Schwierigkeiten auch die dabei notwendigen Rechnungen dem Verständnis darbieten, doch das anzuwendende Verfahren als einen glücklichen Kunstgriff erscheinen lassen und daher zu tieferem Nachdenken herausfordern. Bei Gleichungen, deren Grad den vierten übersteigt, gelang es jedoch trotz mannigfaltigster Versuche und Vorschläge nicht, eine Auflösungsmethode zu finden, obgleich man bald eine Menge spezieller Gleichungen fand, die durch Wurzelgröfsen gelöst werden können. Wie zuerst strenge von Abel und schon vorher, wenn auch nicht ganz einwandfrei, von Ruffini festgestellt wurde, sind die allgemeinsten Gleichungen fünften und höheren Grades nicht durch Wurzelgröfsen lösbar, und bald darauf gelang es Galois, die hinreichenden und notwendigen Bedingungen für die Lösbarkeit der algebraischen Gleichungen durch Wurzelgröfsen zu finden und auf ein Problem der Gruppentheorie zu reduzieren.

Um diese Theorien in diesem letzten Abschnitt zu entwickeln, müssen wir zunächst eine Reihe von Betrachtungen einschalten, aus denen sich diese Resultate auf einfachem und naturgemäfsem Wege ergeben.

---

§ 99. Begriff des algebraischen Körpers.

Jede ganze Funktion einer unbestimmten Veränderlichen x läfst sich nach einer beliebigen ganzen Funktion M(x) vom Grade m als Modul auf eine ganze Funktion von höchstens (m — 1)ten Grade reduzieren (§ 56), die man als ihren Rest nach dem Modul M(x) bezeichnen kann. Während nach einem ganzzahligen Modul nur eine beschränkte Anzahl von Resten existieren, ist die Anzahl der Reste nach einer ganzen Funktion unbeschränkt, da die Koeffizienten beliebige Werte annehmen können. Aber alle Reste lassen sich in diesem Falle linear und homogen durch die m Funktionen

$$1, x, x^2, \ldots x^{m-2}, x^{m-1}$$

darstellen, und es gilt aufserdem der Satz, dafs das Produkt zweier Restfunktionen wieder eine solche ist und daher

ebenfalls so ausgedrückt werden kann. Allgemein kann man zur Darstellung einer Restfunktion andere Gröfsen zu Grunde legen in folgender Weise: Wählt man ein System von n^2 Gröfsen (a_{ik}) $(i, k = 0, 1, \ldots m-1)$ beliebig, jedoch so, dafs die Determinante

$$|a_{ik}| \neq 0 \qquad (i, k = 0, 1, \ldots m-1)$$

wird, so ist es offenbar möglich, jede Funktion des Systems

$$1, x, x^2, \ldots x^{m-2}, x^{m-1}$$

linear und homogen durch die m Funktionen

$$f_i(x) = a_{i0} + a_{i1}x + a_{i2}x^2 + \ldots + a_{i\,m-1}x^{m-1}$$
$$(i = 0, 1, \ldots m-1)$$

darzustellen; und da auch das Umgekehrte gilt, so bilden die m Funktionen

$$f_0(x),\ f_1(x), \ldots f_{m-1}(x)$$

das allgemeinste System, das man zur linearen und homogenen Darstellung jeder ganzen Funktion nach dem Modul $M(x)$ benutzen kann.

Nehmen wir nun noch weiter an, dafs das Modul $M(x)$ eine (im natürlichen Rationalitätsbereiche) irreduktible Funktion ist, die demgemäfs durch $P(x)$ bezeichnet werden mag, so kann man noch einige weitere Folgerungen ziehen. Bilden wir nämlich eine allgemeine Restfunktion

$$u_0 f_0(x) + u_1 f_1(x) + \ldots + u_{m-1} f_{m-1}(x)$$

mit unbestimmten rationalen, nicht sämtlich verschwindenden Koeffizienten $u_0, u_1, \ldots u_{m-1}$, so kann diese, weil sie höchstens vom Grade $m-1$ ist, während $P(x)$ den Grad m hat, nur zu $P(x)$ teilerfremd sein. Es kann daher niemals eine Kongruenz von der Form

$$u_0 f_0(x) + u_1 f_1(x) + \ldots + u_{m-1} f_{m-1}(x) \equiv 0 \mod P(x)$$

bestehen, während andrerseits jede ganze Funktion $f(x)$ einem Ausdruck der linken Seite mit geeigneten Koeffizienten kongruent nach dem Modul $P(x)$ sein mufs. Wenn man ferner eine jede gebrochene Funktion $\frac{f(x)}{g(x)}$ als eine Lösung der Kongruenz

$$g(x)X \equiv f(x) \mod P(x)$$

ansieht, die stets, wenn $g(x)$ nicht durch $P(x)$ teilbar ist, eine einzige Lösung nach dem Modul $P(x)$ besitzt, so gilt für die gebrochene Funktionen bezüglich ihrer Darstellung dasselbe, wie für ganze Funktionen.

Betrachten wir jetzt x als eine durch die Gleichung $P(x) = 0$ definierte algebraische Gröfse m ter Ordnung und alle rationalen Funktionen von x als einen Gattungsbereich, so können wir den soeben dargelegten Sätzen auch die folgende Form geben:

In jedem durch eine algebraische Gröfse x m ter Ordnung gebildeten Gattungsbereich, giebt es stets m Gröfsen $f_0(x), f_1(x), \ldots f_{m-1}(x)$, zwischen denen keine homogene lineare Relation mit nicht sämtlich verschwindenden Koeffizienten besteht, durch die sich aber jede andere algebraische Gröfse des Gattungsbereiches linear und homogen darstellen läfst. Das Produkt zweier Gröfsen des Gattungsbereiches ist wieder im Gattungsbereich enthalten.

Wir wollen nun die Betrachtung umkehren und annehmen, dafs ein System von m Gröfsen $x_1, x_2, \ldots x_m$ gegeben sei, zwischen denen keine lineare und homogene Relation mit rationalen Koeffizienten bestehen kann, wenn diese nicht etwa sämtlich verschwinden, und dafs ferner das Produkt je zweier der Gröfsen $x_1, x_2, \ldots x_m$ linear und homogen durch sie selber darstellbar ist.

Die Gesamtheit aller durch $x_1, x_2, \ldots x_m$ linear darstellbaren Gröfsen wollen wir den durch sie definierten Körper nennen, die so darstellbaren Gröfsen als im Körper enthalten oder ihm zugehörig bezeichnen. Ein System Gröfsen des Körpers, zwischen denen keine lineare homogene Relation mit nicht verschwindenden rationalen Koeffizienten bestehen kann, und durch die jede Gröfse des Körpers aber linear und homogen dargestellt werden kann, nennen wir ein Fundamentalsystem oder eine Basis des Körpers. Die Gröfsen $x_1, x_2, \ldots x_m$ bilden ein solches System, aber man kann aus diesem unzählig viele andere herleiten. Wählt man nämlich m^2 Gröfsen u_{ik} beliebig, jedoch so, dafs die Determinante

$$|u_{ik}| \neq 0 \qquad (i, k = 1, 2, \ldots m)$$

wird, und setzt darauf

$$y_i = \sum_k u_{ik} x_k, \qquad (i, k = 1, 2, \ldots m)$$

so sind die m Gröfsen $y_1, y_2, \ldots y_m$ im Körper enthalten und bilden eine Basis, weil die beiden Modulsysteme $(x_1, x_2, \ldots x_m)$ und $(y_1, y_2 \ldots y_m)$ äquivalent sind. Es ist auch leicht einzusehen, dafs jede Basis des Körpers auf die beschriebene Weise dargestellt werden kann, und dafs vor allem die Anzahl der unabhängigen Gröfsen immer gleich m ist. Diese Zahl m, die somit unabhängig von jeder speziellen Wahl der Basis als die Gesamtanzahl der von einander linear unabhängigen Gröfsen des Körpers charakterisiert ist, nennen wir den **Grad** des Körpers. Das Ziel unserer Betrachtungen besteht nun darin nachzuweisen, dafs die Begriffe Körper und Gattungsbereich nicht von einander verschieden sind. Dazu müssen wir jedoch noch einige Betrachtungen voraufschicken.

Da die Produkte je zweier der Gröfsen $x_1, x_2, \ldots x_m$ im Körper enthalten sein sollen, so bestehen eine Reihe von Gleichungen von der Form

$$x_i x_k = \sum_h a_{ik}^{(h)} x_h, \qquad (h, i, k = 1, 2, \ldots m)$$

wo die Koeffizienten $a_{ik}^{(h)}$ rationale Zahlen sind. Mit Hülfe dieser Gleichungen ist es möglich, alle ganzen Funktionen von $x_1, x_2, \ldots x_m$ linear durch $x_1, x_2, \ldots x_m$ auszudrücken. Dasselbe gilt aber auch für jede gebrochene Funktion $\frac{y}{x}$, wenn nicht der Nenner x verschwindet. Da nämlich zwischen den Produkten

$$x x_1, x x_2, \ldots x x_m$$

keine lineare und homogene Relation besteht, weil aus ihr eine solche durch einfache Division mit x zwischen $x_1, x_2, \ldots x_m$ ableitbar wäre, so bilden diese ein Fundamentalsystem, es läfst sich also y durch dieses linear und homogen darstellen, und hieraus folgt dann für $\frac{y}{x}$ eine lineare und homogene Darstellung in $x_1, x_2, \ldots x_m$. Nehmen wir $y = x$ an, so ergiebt sich, dafs die Einheit und somit alle rationalen Zahlen im Körper enthalten sind. Fassen wir alles zusammen, so erhalten wir den Satz:

Jeder Körper enthält alle rationalen Zahlen und alle rationalen Funktionen der in ihm enthaltenen Gröfsen.

Bilden wir nun von einer im Körper enthaltenen Gröfse

$$x = u_1 x_1 + u_2 x_2 + \ldots + u_m x_m$$

die Reihe der Potenzen

$$1, x, x^2, x^3, \ldots,$$

so können unter diesen nicht mehr als m unabhängige vorkommen. Nehmen wir an, dafs zwischen

$$1, x, x^2, \ldots x^{n-1}$$

keine lineare und homogene Relation mit rationalen Koeffizienten bestehen kann, während x^n in der Form

$$F(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots + a_n = 0$$

sich durch $1, x, x^2, \ldots x^{n-1}$ ausdrücken lasse, so gilt dasselbe umsomehr von den höheren Potenzen, $F(x)$ ist eine irreduktible Gleichung, der die Gröfse x genügt. Es ergiebt sich:

Jede Gröfse x des Körpers ist eine algebraische Gröfse. Nach dem Grad der irreduktiblen Gleichung, der sie genügt, bezeichnet man die Gröfse x als eine algebraische Gröfse n ter Ordnung. Aus dieser Gröfse läfst sich ein Gattungsbereich bilden, dessen Gröfsen alle im Körper enthalten sind. Könnten wir nun zeigen, dafs es im Körper algebraische Gröfsen m ter Ordnung giebt, so wäre die Identität der Begriffe Gattungsbereich und Körper dargethan. Bevor wir dies ausführen, müssen wir zunächst untersuchen, wie sich die Grade zweier Körper K und H zu einander verhalten, von denen die Gröfsen des zweiten alle im ersten enthalten sind. Man sagt dann oft kurz, dafs der Körper K den Körper H enthalte, oder dafs H ein Teil von K sei.

Nehmen wir an, dafs der Körper H vom Grade n sei, und dafs die Gröfsen

$$y_0, y_1, \ldots y_{n-1}$$

ein zu ihm gehöriges Fundamentalsystem darstellen. Nehmen wir dann eine Gröfse z hinzu und bilden die Produkte

$$y_0 z, y_1 z, \ldots y_{n-1} z,$$

so sind diese entweder alle in H enthalten, oder es kann zwischen ihnen und den Gröfsen $y_0, y_1, \ldots y_{n-1}$ von H keine lineare Beziehung existieren, je nachdem z in H enthalten ist oder nicht. Eine lineare homogene Relation würde dazu dienen können, z als gebrochene lineare Funktion von $y_0, y_1, \ldots y_{n-1}$, und somit als eine Gröfse des Körpers H darzustellen. Betrachten wir allgemeiner die Reihen

$$\begin{array}{llll} y_0, & y_1, & \ldots\ldots\ldots & y_{n-1} \\ y_0 z_1, & y_1 z_1, & \ldots\ldots & y_{n-1} z_1 \\ y_0 z_2, & y_1 z_2, & \ldots\ldots & y_{n-1} z_2 \\ \vdots & & & \\ y_0 z_{i-1}, & y_1 z_{i-1}, & \ldots & y_{n-1} z_{i-1} \end{array}$$

und nehmen an, dafs zwischen den Elementen keine lineare und homogene Relation mit rationalen Koeffizienten bestehen kann, so liefert ein neues Element z_i eine Reihe

$$y_0 z_i, y_1 z_i, \ldots y_{n-1} z_i$$

von n neuen Elementen, die entweder alle durch die voraufgegangenen linear darstellbar sind, oder zwischen denen überhaupt keine lineare Relation bestehen kann. Wenn nämlich eine lineare Relation besteht, so gestattet diese z_i und somit $y_k z_i$ als gebrochene lineare Funktion der vorhergehenden Elemente darzustellen, wo der Nenner aber eine Gröfse des Körpers H ist, so dafs die gebrochene Funktion in eine ganze umgeformt werden kann.

Wenden wir dies nun an auf den Körper K, den wir als verschieden von H voraussetzen. Dann mufs es in K eine Gröfse z geben, die nicht in H enthalten ist, und es müssen dann die Gröfsen

$$y_0 z_1, y_1 z_1, \ldots y_{n-1} z_1$$

n neue linear unabhängige Elemente von K sein. Enthält K nun noch eine weitere Gröfse z_2, die durch die Gröfsen y_i und $y_i z_1$ nicht linear dargestellt werden kann, so liefert uns die Reihe

$$y_0 z_2, y_1 z_2, \ldots y_{n-1} z_2$$

wieder n neue Elemente von K, zwischen denen und den schon genannten keine lineare Abhängigkeit existiert. So kann man weiter fortfahren, wenn noch kein Fundamental-

system von K gebildet ist, mufs aber endlich auf ein Element z_{r-1} und damit auf eine Reihe

$$y_0 z_{r-1}, \; y_1 z_{r-1}, \ldots y_{n-1} z_{r-1}$$

stofsen, so dafs zwischen den Elementen dieser und denen der vorhergehenden Reihen keine lineare Bezeichnung obwaltet, aber jede Gröfse von K linear durch sie dargestellt werden kann. Dann bilden die n r Produkte

$$y_i, \; y_i z_k$$
$$(i = 0, 1, \ldots n-1; \; k = 1, \ldots r-1)$$

ein Fundamentalsystem des Körpers K vom Grade $m = rs$.

Wir gewinnen also durch diese Betrachtung den Satz:

Enthält ein Körper K einen andern H, so ist sein Grad m ein Vielfaches des Grades n von H.

Einen Körper, der keine andern Körper enthalten kann, nennt man einen *Primkörper*. *Ist der Grad eines Körpers eine Primzahl p, so ist er also sicher ein Primkörper*, denn dann kann er nur *Körper ersten Grades* enthalten, wenn man diese Bezeichnung in uneigentlicher Bedeutung für die Gesamtheit aller rationalen Zahlen gebrauchen will.

Alle Gröfsen des Körpers K lassen sich linear und homogen durch $\frac{m}{n} = r$ Elemente $1, z_1, z_2, \ldots z_{r-1}$ darstellen, wenn man als Koeffizienten nicht nur rationale Zahlen, sondern Gröfsen des Körpers H, also lineare Formen von $y_0, y_1, \ldots y_{n-1}$ mit rationalen Koeffizienten zuläfst, und zwischen diesen Elementen besteht ebenfalls keine lineare und homogene Relation mit Koeffizienten, die dem Körper H angehören. Daher kann man die Gröfsen $1, z_1, z_2, \ldots z_{r-1}$ als Fundamentalsystem des Körpers K in Bezug auf H ansehen, wobei der Körper H dieselbe Rolle spielt, wie früher die Gesamtheit der rationalen Zahlen, die man ja ebenfalls als einen uneigentlichen Körper ansehen kann. Wenn man so vom Körper H absehen will, so drückt man dies häufig so aus, dafs man sagt, dieser Körper werde dem Rationalitätsbereich *adjungiert*. Wir erhalten somit den Satz:

Enthält der Körper K vom Grade m einen andern H vom Grade n, so ist K ein relativer

Körper vom Grade $\frac{m}{n} = r$, wenn man den Körper H dem Rationalitätsbereich adjungiert.

Um nun den Nachweis zu führen, dafs in jedem Körper K vom Grade m auch algebraische Gröfsen m ten Grades, die man auch **primitive Gröfsen** des Körpers nennt, vorhanden sind, schlagen wir folgenden Weg ein.

Wir nehmen an, dafs ein Körper H vom Grade n vorliege, in dem y eine primitive Gröfse sei, also

$$1,\ y,\ y^2, \ldots y^{n-1}$$

ein Fundamentalsystem bilden. Wir nehmen nun irgend eine nicht im Körper H, aber in K vorkommende Gröfse z hinzu und bilden die Reihen

$$\begin{array}{llll} z, & y\,z, & y^2 z, & \ldots\ldots y^{n-1} z \\ z^2, & y\,z, & y^2 z^2, & \ldots\ldots y^{n-1} z^2 \\ \vdots & & & \\ z^{r-1}, & y\,z^{r-1}, & y^2 z^{r-1}, & \ldots y^{n-1} z^{r-1}, \end{array}$$

bis wir zum ersten Male auf eine Potenz z^r stofsen, die sich linear und homogen durch die Elemente der vorhergehenden Reihen darstellen läfst. Wie leicht zu erkennen ist, bilden dann diese Elemente einen Körper H′ vom Grade $n' = n\,r$, und wir wollen nun nachweisen, dafs H′ eine primitive Gröfse x vom Grade n′ besitzt. Verstehen wir unter u und v zwei unbestimmte rationale Zahlen, und betrachten wir die Gröfse

$$x = u\,y + v\,z,$$

so wird diese als algebraische Gröfse in H′ einer irreduktibeln Gleichung

$$F(x;\ u, v) = 0$$

genügen, deren Koeffizienten rational sind. Da diese Gleichung für jeden Wert von u und v besteht, so kann man sie nach u und v differenzieren. Hierbei findet man

$$F'(x)\,y + \frac{\partial F}{\partial u} = 0, \quad F'(x)\,z + \frac{\partial F}{\partial v} = 0.$$

Weil nun $F(x;\ u, v)$ eine irreduktible Funktion ist, so ist sie zu $F'(x;\ u, v)$ teilerfremd; es existiert also eine Gleichung von der Form

$$F(x;\ u, v)\,G_1(x;\ u, v) + F'(x;\ u, v)\,G(x;\ u, v) = R(u, v),$$

wo alle Funktionen ganz und rational von den Argumenten abhängen, und $R(u, v)$ nicht für jeden Wert von u und v verschwinden kann. Verfügen wir nun über u und v so, dafs dies nicht geschieht, so wird auch $F'(x; u, v) \neq 0$, und es folgt daher durch Division damit

$$y = -\frac{\frac{\partial F}{\partial u}}{F'(x)}, \quad z = -\frac{\frac{\partial F}{\partial v}}{F'(x)}.$$

Also sind y und z rational durch eine einzige algebraische Gröfse x darstellbar, die ihrerseits wieder rational und zwar linear von y und z abhängt. Daher ist x eine primitive Gröfse des Körpers H'. Verfahren wir nun, wenn H' noch nicht mit K übereinstimmt, ebenso mit H' wie vorher mit H, so finden wir eine primitive Gröfse eines Körpers H'', der H' enthält aber von K ein Teil ist u. s. w. Man erhält so von einer Reihe von Körpern $H, H', H'', H''', \ldots$ primitive Gröfsen. Da die Grade $n, n', n'', n''', \ldots$ fortwährend zunehmen, so findet das Verfahren einen natürlichen Abschlufs dadurch, dafs der letzte Körper $H^{(s)}$ mit K übereinstimmt. Damit sind wir am Ziele und haben den Satz:

In jedem Körper $m$ten Grades giebt es algebraische Gröfsen $m$ter Ordnung oder primitive Elemente.

Wir haben alle diese Untersuchungen unter der Voraussetzung durchgeführt, dafs kein algebraischer Körper dem Rationalitätsbereich adjungiert worden ist. Diese Voraussetzung ist aber durchaus unwesentlich, und alle Betrachtungen lassen sich Wort für Wort auf den Fall verallgemeinern, dafs nicht mehr ein natürlicher Rationalitätsbereich, sondern ein durch Adjunktion eines Körpers entstehender Gattungsbereich vorliegt. Es bleiben somit auch dann alle entwickelten Sätze für solche Körper von Bestand.

---

§ 100. Symmetrische Funktionen.

Am Schlusse des vorigen Abschnittes haben wir erkannt, dafs jede ganze Funktion in lauter Linearfaktoren zerlegbar ist, wenn man irrationale und komplexe Wurzeln einführt.

Ist

$$F(x) = x^n - c_1 x^{n-1} + c_2 x^{n-2} + \ldots + (-1)^n c_n$$

eine ganze Funktion, und sind $x_1, x_2, \ldots x_n$ die Wurzeln der Gleichung $F(x) = 0$, so kann man

$$F(x) = (x - x_1)(x - x_2) \ldots (x - x_n)$$

setzen, und wenn man nun die Koeffizienten gleicher Potenzen von x in diesen beiden Darstellungen vergleicht, so ergiebt sich

$$\begin{aligned} f_1(x_1, x_2, \ldots x_n) &= c_1 \\ f_2(x_1, x_2, \ldots x_n) &= c_2 \\ &\vdots \\ f_n(x_1, x_2, \ldots x_n) &= c_n, \end{aligned}$$

wo $f_1, f_2, \ldots f_n$ die in § 5 erwähnten n symmetrischen Grundformen der n Wurzeln $x_1, x_2, \ldots x_n$ unserer Gleichung sind. Da $c_1, c_2, \ldots c_n$ gegebene Gröfsen sind, so ergiebt sich hieraus, dafs die symmetrischen Grundformen der Wurzeln jederzeit bekannt sind. Diese Thatsache kann nun bedeutend verallgemeinert werden auf beliebige symmetrische Funktionen der Wurzeln. Hierzu sind aber vorerst einige allgemeine Betrachtungen erforderlich, zu deren Darlegung wir jetzt schreiten.

Unter einer symmetrischen Funktion von mehreren Veränderlichen $x_1, x_2, \ldots x_n$ versteht man eine solche Funktion, die ihre Form beibehält, wenn man die Veränderlichen beliebig untereinander vertauscht. Man kann auch sagen, dafs sie ihren Wert nicht ändert, vorausgesetzt, dafs man die Variabeln als völlig unbestimmt ansieht, also ihnen beliebige Werte erteilen darf. Denn es kann sehr wohl vorkommen, dafs eine nicht symmetrische Funktion ihren Wert bei allen Vertauschungen beibehält, wenn den Variabeln besondere Werte beigelegt werden. Da alle Permutationen von n Gröfsen sich durch Zusammensetzung von $(n-1)$ Transpositionen mit einem gemeinsamen Element erzeugen lassen, so genügt es, die auf ihre Symmetrie zu untersuchende Funktion den Transpositionen

$$(x_1, x_2),\ (x_1, x_3), \ldots (x_1, x_n)$$

zu unterwerfen. Dafs sie dadurch nicht geändert wird, ist die hinreichende und notwendige Bedingung dafür, dafs sie symmetrisch ist.

Summe, Differenz, Produkte und Quotienten symmetrischer Funktionen sind selbst wieder symmetrisch. Bei der Betrachtung der Gesamtheit symmetrischer Funktionen können wir uns auf ganze Funktionen beschränken. Denn jede gebrochene symmetrische Funktion läfst sich als Quotient zweier ganzen symmetrischen Funktionen darstellen. Dies ist allerdings nicht immer der Fall, wenn man die Funktion als Quotient zweier ganzen Funktionen darstellt. Vertauscht man aber in einer derselben, etwa im Nenner, die Variabeln auf alle mögliche Arten, so wird man aus ihr eine Reihe neuer Funktionen gewinnen. Erweitert man sodann den Quotienten mit dem Produkt aller so erhaltenen verschiedenen Funktionen, so wird der Nenner eine symmetrische Funktion, und da der Quotient symmetrisch ist, mufs es der neue Zähler ebenfalls sein.

Kommt in einer ganzen symmetrischen Funktion von $x_1, x_2, \dots x_n$ das Glied $A x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ vor, so kommen in ihr auch alle die Glieder vor, die man aus diesem durch Vertauschung von $x_1, x_2, \dots x_n$ herleiten kann. Man erhält so aus einem einzigen Gliede eine homogene symmetrische Funktion. Jede ganze symmetrische Funktion läfst sich also als Aggregat von lauter symmetrischen Formen darstellen. Diese sind vollständig durch das Exponentensystem $(a_1, a_2, \dots a_n)$ charakterisiert und werden auch als Vandermondesche Typen bezeichnet, und es ist offenbar wegen der Symmetrie gestattet, anzunehmen, dafs

$$a_1 \geqq a_2, \quad a_2 \geqq a_3, \dots a_{n-1} \geqq a_n$$

ist.

Zu den symmetrischen Grundformen gelangen wir nun auf einem andern Wege wie früher, wenn wir die Typen in Betracht ziehen, die die Variabeln nur in der ersten Potenz enthalten. Es ist klar, dafs dann nur n Exponentensysteme

$$(1, 0, \dots 0), \ (1, 1, 0, \dots 0), \dots (1, 1, \dots 1)$$

möglich sind, denen n Formen vom 1 ten, 2 ten, ... n ten Grade entsprechen. Wenn man von konstanten Faktoren absieht, die diese Formen besitzen können, so ersieht man sofort, dafs sie in der That mit den symmetrischen Grundformen übereinstimmen.

Aus dieser neuen Definition der symmetrischen Grundformen wird nun aber sofort klar, dafs man mit ihrer Hülfe symmetrische Formen konstruieren kann, in denen das Glied

$$A x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$$

das Glied von höchster Ordnung ist, wenn wir die Glieder der symmetrischen Form in der Weise anordnen, wie das in § 3 auseinandergesetzt ist. Bilden wir nämlich die symmetrische Form

$$\Phi(x_1, x_2, \dots x_n) = A f_1^{a_1 - a_2}(x_1, \dots x_n) f_2^{a_2 - a_3}(x_1, \dots x_n) \dots f_{n-1}^{a_{n-1} - a_n}(x_1, \dots x_n) f_n^{a_n}(x_1, \dots x_n),$$

so sind die Glieder höchster Ordnung in $f_1^{a_1 - a_2}$, $f_2^{a_2 - a_3}$, $\dots f_{n-1}^{a_{n-1} - a_n}$, $f_n^{a_n}$, resp. $x_1^{a_1 - a_2}$, $(x_1 x_2)^{a_2 - a_3}, \dots (x_1 x_2 \dots x_{n-1})^{a_{n-1} - a_n}$, $(x_1 x_2 \dots x_n)^{a_n}$ und daher ist das Glied höchster Ordnung in dem Produkt

$$x_1^{a_1 - a_2}(x_1 x_2)^{a_2 - a_3} \dots (x_1 x_2 \dots x_{n-1})^{a_{n-1} - a_n}(x_1 \dots x_n)^{a_n} = x_1^{a_1} x_2^{a_2}, \dots x^{a_n}$$

wie behauptet wurde. Dafs dieses Potenzprodukt auch den Koeffizienten A hat, ist dann sofort ersichtlich, da die Koeffizienten in den symmetrischen Grundformen lauter Einheiten sind.

Ist nun $F(x_1, x_2, \dots x_n)$ eine beliebige symmetrische Funktion und $A x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ das Glied Glied höchster Ordnung in ihr, so ist die Differenz

$$F(x_1, x_2, \dots x_n) - \Phi(x_1, x_2, \dots x_n) = F_1(x_1, x_2, \dots x_n)$$

ebenfalls eine symmetrische Funktion, aber eine solche, in der das Glied höchster Ordnung niedriger ist als in $F(x_1, x_2, \dots x_n)$. Auf diese Funktion können wir dasselbe Schlufsverfahren anwenden wie auf F, also eine aus den symmetrischen Grundformen gebildete Form $\Phi_1(x_1, x_2, \dots x_n)$ so bestimmen, dafs in der Differenz

$$F_1(x_1, x_2, \dots x_n) - \Phi_1(x_1, x_2, \dots x_n) = F_2(x_1, x_2, \dots x_n)$$

das Glied höchster Ordnung niedriger ist als bei F_1. Da die Ordnung des höchsten Gliedes durch diese Art der Reduktion beständig erniedrigt wird, und es nur eine beschränkte Anzahl von Gliedern niederer Ordnung giebt, so mufs das

Verfahren schliefslich von selbst abbrechen, indem zuletzt ein Glied F_m zum Vorschein kommt, das mit gar keinen Variabeln behaftet ist. Erhält man bei dieser Reduktion die Reihe von Gleichungen

$$\begin{aligned} F &= \Phi + F_1 \\ F_1 &= \Phi_1 + F_2 \\ &\vdots \\ F_{m-1} &= \Phi_{m-1} + F_m, \end{aligned}$$

so ergiebt sich aus diesen die Darstellnng

$$F = \Phi + \Phi_1 + \Phi_2 + \ldots + \Phi_{m-1} + F_m$$

und damit der Satz:

Jede ganze symmetrische Funktion der n Variabeln $x_1, x_2, \ldots x_n$ läfst sich als ganze Funktion der n symmetrischen Grundformen

$$f_1(x_1, x_2, \ldots x_n), f_2(x_1, x_2, \ldots x_n), \ldots f_n(x_1, x_2, \ldots x_n)$$

darstellen.

Beachten wir nun noch die obigen Bemerkungen über gebrochene rationale symmetrische Funktionen, so erhalten wir weiter:

Jede rationale symmetrische Funktion der n Variabeln $x_1, x_2, \ldots x_n$ ist als rationale Funktion der symmetrischen Grundformen darstellbar.

Zu Anfang dieses Paragraphen hatten wir uns die Variabeln $x_1, x_2, \ldots x_n$ als Wurzeln einer algebraischen Gleichung vorgestellt und gesehen, dafs die symmetrischen Grundformen in sehr einfacher Weise mit den Koeffizienten zusammenhängen. Indem wir die Untersuchung hierauf spezialisieren, finden wir:

Jede rationale symmetrische Funktion aller Wurzeln einer algebraischen Gleichung ist eine rationale Funktion ihrer Koeffizienten.

Obgleich in den vorhergehenden Entwickelungen auch eine allgemeine Methode der Darstellung jeder symmetrischen Funktion durch die symmetrischen Grundformen enthalten ist, so läfst sich die Darstellung unter Umständen noch in besonders einfacher Weise herleiten. Wir wollen dies für die speziellen symmetrischen Funktionen

$$s_k = x_1^k + x_2^k + \ldots + x_n^k \qquad (k = 0, 1, 2 \ldots)$$

die man als **Potenzsummen der Wurzeln** bezeichnet, näher darlegen. Man erkennt sofort, dafs

$$s_0 = n, \quad s_1 = c_1$$

ist, und aus

$$x_i^{k-n} F(x_i) = x_i^k - c_1 x_i^{k-1} + \ldots + (-1)^{n-1} c_{n-1} x_i^{k-n+1} + (-1)^n c_n x_i^{k-n} = 0 \quad (i = 1, 2, \ldots n;\ k \geqq n)$$

ergiebt sich durch Summation sofort die allgemeine für $k \geqq n$ geltende Rekursionsformel

$$s_k - c_1 s_{k-1} + c_2 s_{k-2} - \ldots + (-1)^{n-1} c_{n-1} s_{k-n+1} + (-1)^n c_n s_{k-n} = 0,$$

die $s_n, s_{n+1}, s_{n+2}, \ldots$ u. s. w. berechnen lehrt, wenn $s_0, s_1, \ldots s_{n-1}$ bekannt sind. Um diese letzteren zu finden, bilden wir

$$\frac{F(x) - F(x_i)}{x - x_i} = x^{n-1} + f_1(x_i) x^{n-2} + \ldots + f_{n-2}(x_i) x + f_{n-1}(x_i),$$

wobei

$$f_1(x_i) = x_i - c_1,$$
$$f_2(x_i) = x_i^2 - c_1 x_i + c_2,$$
$$\vdots$$
$$f_{n-1}(x_1) = x_i^{n-1} - c_1 x_i^{n-2} + \ldots + (-1)^{n-1} c_{n-1}$$

gesetzt ist. (Vgl. § 59). Nun ergiebt sich aber

$$\sum_i \frac{F(x)}{x - x_i} = F'(x) = n x^{n-1} - (n-1) c_1 x^{n-2} + \ldots + (-1)^{n-1} c_{n-1}, \qquad (i = 1, 2, \ldots n)$$

während andererseits sich durch Summation

$$n x^{n-1} + \sum_i f_1(x_i) x^{n-2} + \ldots + \sum_i f_{n-2}(x_i) x + \sum_i f_{n-1}(x_i) \qquad (i = 1, 2, \ldots n)$$

folgt. Vergleicht man die Koeffizienten gleich hoher Potenzen so erhält man

$$\sum_i f_k(x_i) = (-1)^k (n-k) c_k \qquad (i = 1, 2, \ldots n;\ k = 1, 2, \ldots n-1)$$

und daraus die Rekursionsformeln für $s_1, s_2, \ldots s_{n-1}$ in der Gestalt

$$0 = s_1 - c_1$$
$$0 = s_2 - c_1 s_1 + 2 c_2$$
$$\vdots$$
$$0 = s_{n-1} - c_1 s_{n-2} + c_2 s_{n-3} - \ldots + (-1)^{n-1} (n-1) c_{n-1}.$$

Hieraus folgt z. B.

$$s_1 = c_1$$
$$s_2 = c_1^2 - 2 c_2$$
$$s_3 = c_1^3 - 3 c_1 c_2 + 3 c_3$$
$$s_4 = c_1^4 - 4 c_1^2 c_2 + 4 c_1 c_3 + 2 c_2^2 - 4 c_4$$

u. s. w. Damit ist eine einfache Berechnung aller Potenzsummen dargelegt.

Wir haben im neunten Abschnitt in § 77 und § 79 auf die einfache Gestalt aufmerksam gemacht, auf die die Resultante und die Discriminante gebracht werden können, wenn die Funktionen in Linearfaktoren zerlegt sind. Nachdem wir im vorigen Abschnitte gesehen haben, dafs eine solche Zerlegung stets möglich ist, wenn man zur Einführung irrationaler und komplexer Wurzeln seine Zuflucht nimmt, ist es jetzt am Platze, darauf hinzuweisen, dafs die Resultante und Discriminante symmetrische Funktionen der Wurzeln sind und als solche nach der in diesem Paragraphen entwickelten Methode rational durch die Koeffizienten dargestellt werden können. Wir erhalten so ein neues Verfahren, sie zu berechnen, das sich als sehr praktisch erweist.

---

§ 101. Reduktion ganzer rationaler Funktionen der Wurzeln.

Die Anwendung der Theorie der symmetrischen Funktionen auf die Wurzeln algebraischer Gleichungen gründet sich nach den vorhergehenden Entwickelungen auf die Existenz der Wurzeln, die wir in § 93 bewiesen hatten. Statt aber in dieser Weise mit Einführung irrationaler und complexer Gröfsen vorzugehen, kann man auch die Sache von einem rein arithmetischen Standpunkte aus betrachten, wenn man sich der Theorie der Modulsysteme bedient. Vereinigen wir nämlich, wenn wie im vorigen Paragraphen

$$F(x) = x^n - c_1 x^{n-1} + c_2 x^{n-2} - \ldots \pm c_n$$

die zu Grunde liegende ganze Funktion ist, die Gröfsen

$$f_1(x_1, \ldots x_n) - c_1,\ f_2(x_1, \ldots x_n) - c_2, \ldots f_n(x_1, \ldots x_n) - c_n$$

zu einem Modulsystem, so läfst sich zunächst behaupten, dafs nach diesem

$$F(x) \equiv (x - x_1)(x - x_2) \ldots (x - x_n),$$

also in ein Produkt von lauter Linearfaktoren zerlegbar ist. Dies geht nämlich unmittelbar aus der leicht zu bestätigenden identischen Gleichung

$$F(x) = (x - x_1)(x - x_2) \ldots (x - x_n) + (f_1 - c_1) x^{n-1}$$
$$- (f_2 - c_2) x^{n-2} + \ldots - (-1)^n (f_n - c_n)$$

hervor, die $F(x)$ linear durch das Produkt $(x - x_1)(x - x_2) \ldots (x - x_n)$ und die Elemente des Modulsystems darstellen lehrt. Eine solche Auffassung enthält aber auch die bisherige in sich; denn wenn $x_1, x_2, \ldots x_n$ die Wurzeln der Gleichung $F(x) = 0$ sind, so verschwinden, wie die Gleichungen am Anfang des vorigen Paragraphen zeigen, die Elemente des Modulsystems sämtlich, und es geht daher die obige Kongruenz in eine Gleichung über.

Wenn wir nun dieses Modulsystem einer näheren Betrachtung unterziehen, so ergeben sich die im vorigen Paragraphen entwickelten Sätze nicht nur von einem allgemeinern Gesichtspunkte aus, sondern wir gelangen auch dazu, jede beliebige rationale Funktion von $x_1, x_2, \ldots x_n$ zu reduzieren.

Wir transformieren zunächst unser Modulsystem so, dafs an Stelle der symmetrischen Grundformen der n Gröfsen $x_1, x_2, \ldots x_n$ solche der $(n - 1)$ Gröfsen $x_2, x_3, \ldots x_n$ auftreten. Zunächst ist nämlich

$$f_1(x_1, x_2, \ldots x_n) - c_1 = f_1(x_2, x_3, \ldots x_n) - (c_1 - x_1)$$

oder, wenn

$$y_1 = -x_1 + c_1$$

gesetzt wird,

$$f_1(x_1, x_2, \ldots x_n) = f_1(x_2, x_3, \ldots x_n) - y_1.$$

Weiter ergiebt sich

$$f_2(x_1, x_2, \ldots x_n) - c_2 = f_2(x_2, x_3, \ldots x_n) + x_1 f_1(x_2, x_3, \ldots x_n) - c_2$$
$$= f_2(x_2, x_3, \ldots x_n) + x_1 [f_1(x_2, x_3, \ldots x_n) - y_1] + y_1 x_1 - c_2$$

und somit bei Einführung der Gröfse

$$y_2 = - y_1 x_1 + c_2$$

einfacher

$$f_2(x_1, x_2, \dots x_n) - c_2 = f_2(x_2, x_3, \dots x_n) - y_2 + x_1 [f_1(x_2, x_3, \dots x_n) - y_1].$$

Aus den beiden Gleichungen folgt nun aber die Äquivalenz

$$[f_1(x_1, x_2, \dots x_n) - c_1, f_2(x_1, x_2, \dots x_n) - c_2] = [f_1(x_2, x_3, \dots x_n) - y_1, f_2(x_2, x_3, \dots x_n) - y_2].$$

Diese Äquivalenz kann man nun leicht verallgemeinern und zeigen, dafs

$$[f_1(x_1, x_2, \dots x_n) - c_1, f_2(x_1, x_2, \dots x_n) - c_2, \dots f_i(x_1, x_2, \dots x_n) - c_i] = [f_1(x_2, x_3, \dots x_n) - y_1, f_2(x_2, x_3, \dots x_n) - y_2, \dots f_i(x_2, x_3, \dots x_n) - y_i] \qquad (i = 1, 2, \dots n)$$

wird, wenn allgemein

$$y_k = - y_{k-1} x_1 + c_k \qquad (k = 1, 2, \dots n)$$

gesetzt wird. Denn nehmen wir an, dafs die Äquivalenz für einen gewissen Wert von i schon bewiesen ist, so geht aus der Transformation

$$f_{i+1}(x_1, x_2, \dots x_n) - c_{i+1} = f_{i+1}(x_2, x_3, \dots x_n) + x_1 f_i(x_2, x_3, \dots x_n) - c_{i+1}$$
$$= f_{i+1}(x_2, x_3, \dots x_n) - y_{i+1} + x_1 [f_i(x_2, x_3, \dots x_n) - y_i]$$

sofort hervor, dafs die Äquivalenz für den Wert $(i + 1)$ ebenfalls gilt. Die Funktionen $y_1, y_2, \dots y_n$ lassen sich sehr leicht bestimmen. Man erhält

$$y_1 = -(x_1 - c_1), \; y_2 = x_1^2 - c_1 x_1 + c_2,$$
$$y_3 = -(x_1^3 - c_1 x_1^2 + c_2 x_1 - c_3)$$

und kann leicht durch vollständige Induktion zeigen, dafs allgemein

$$y_i = (-1)^i [x_1^i - c_1 x_1^{i-1} + c_2 x_1^{i-2} - \dots + (-1)^{i-1} c_{i-1} x_1 + (-1)^i c_i]$$

ist, also speziell für $i = n$

$$y_n = (-1)^n F(x_1)$$

wird. Unser Modulsystem ist also transformiert in

$$[f_1(x_2, x_3, \dots x_n) - y_1, f_2(x_2, x_3, \dots x_n) - y_2, \\ \dots f_{n-1}(x_2, x_3, \dots x_n) - y_{n-1}, F(x_1)].$$

Auf das in diesem System enthaltene Teilsystem

$$[f_1(x_2, x_3, \dots x_n) - y_1, f_2(x_2, x_3, \dots x_n) - y_2, \\ \dots f_{n-1}(x_2, x_3, \dots x_n) - y_{n-1}]$$

können wir nun dieselbe Art der Transformation anwenden, wie beim ursprünglichen, es also überführen in

$$[f_1(x_3, x_4, \dots x_n) - y_1', f_2(x_3, x_4, \dots x_n) - y_2', \\ \dots f_{n-2}(x_3, x_4, \dots x_n) - y'_{n-2}, F_1(x_1, x_2)],$$

und hierbei ist

$$y_i' = (-1)^i \sum_{h=0}^{h=i} (-1)^h y_h x_2^{i-h}$$

oder

$$y_i' = (-1)^i \sum_{(h, k_1, k_2)} (-1)^h c_h x_1^{k_1} x_2^{k_2},$$

$$(h, k_1, k_2 = 0, 1, \dots i;\ h + k_1 + k_2 = i)$$

speziell

$$y'_{n-1} = (-1)^{n-1} F_1(x_1, x_2),\ F_1(x_1, x_2) = \sum_{(h, k_1, k_2)} (-1)^h c_h x_1^{k_1} x_2^{k_2}$$

$$(h, k_1, k_2 = 0, 1, \dots n-1;\ h + k_1 + k_2 = n-1)$$

gesetzt worden.

So kann man weiter fortfahren. Man gewinnt allgemein ein Modulsystem von folgender Gestalt

$$[f_1(x_{r+1}, \dots x_n) - y_1^{(r-1)}, \dots f_{n-r}(x_{r+1}, \dots x_n) - y_{n-r}^{(r-1)}, \\ F_{r-1}(x_1, x_2, \dots x_r), F_{r-2}(x_1, x_2, \dots x_{r-1}), \dots F_1(x_1, x_2), F(x_1)]$$

und hierbei ist

$$y_i^{(r-1)} = (-1)^{i-1} \sum_{(h, k_1, k_2, \dots k_r)} (-1)^h c_h x_1^{k_1} x_2^{k_2} \dots x_r^{k_r},$$

$$(h, k_1, k_2, \dots k_r = 0, 1, \dots i;\ h + k_1 + k_2 + \dots + k_r = i)$$

speziell

$$F_{r-1}(x_1, x_2, \dots x_r) = \sum_{(h, k_1, k_2, \dots k_r)} (-1)^h c_h x_1^{k_1} x_2^{k_2} \dots x_r^{k_r}.$$

$$(h, k_1, k_2, \dots k_r = 0, 1, \dots n-r+1;\ h + k_1 + k_2 + \dots + k_r = n-r+1)$$

Als Schlufsresultat gewinnen wir somit die folgende Äquivalenz:

$$[f_1(x_1, x_2, \ldots x_n) - c_1, f_2(x_1, x_2, \ldots x_n) - c_2, \ldots f_n(x_1, x_2, \ldots x_n) - c_n]$$
$$= [F(x_1), F_1(x_1, x_2), \ldots F_{n-2}(x_1, x_2, \ldots x_{n-1}), F_{n-1}(x_1, x_2, \ldots x_n)],$$

wobei bemerkt werden mag, dafs sich $F_{n-1}(x_1, x_2, \ldots x_n)$ nicht von $f_1(x_1, x_2, \ldots x_n) - c_1$ unterscheidet.

Mit Hülfe des auf der rechten Seite stehenden Modulsystems kann man nun jede beliebige ganze Funktion der n Variabeln $x_1, x_2, \ldots x_n$ auf eine besonders einfache Form reduzieren. Man kann zunächst $F_{n-1}, (x_1, x_2, \ldots x_n)$ dazu benutzen, um die Funktion von x_n zu befreien. Darauf dient $F_{n-2}(x_1, x_2, \ldots x_{n-1})$ dazu, um aus der Funktion alle höheren Potenzen von x_{n-1} fortzuschaffen. Mit $F_{k-1}(x_1, x_2, \ldots x_k)$ kann man allgemein die $(n - k + 1)$ten und höheren Potenzen von x_k entfernen. Schliefslich erlaubt es $F(x_1)$, zu erzielen, dafs die Funktion keine höhere Potenz von x_1 als die $(n-1)$te enthält. Fassen wir alles zusammen, so läfst sich jede ganze Funktion von $x_1, x_2, \ldots x_n$ auf die Form reduzieren

$$\sum_{(i_1, i_2, \ldots i_{n-1})} A_{i_1 i_2 \ldots i_{n-1}} x_1^{i_1} x_2^{i_2} \ldots x_{n-1}^{i_{n-1}},$$

wo

$$0 \leqq i_k \leqq n - k \qquad (k = 1, 2, \ldots n-1)$$

ist.

Aus dieser Art der Darstellung läfst sich nun auch die Richtigkeit des im vorigen Paragraphen bewiesenen Satzes ableiten, dafs jede symmetrische Funktion rational durch die Grundformen ausgedrückt werden, wenn man dabei benutzt, dafs eine ganze Funktion $f(x)$ $m$ten Grades nicht für $(m + 1)$ verschiedene Werte der Variabeln denselben Wert a annehmen kann. Dies ergiebt sich leicht daraus, dafs dann $f(x) - a$ für mehr als m Werte verschwinden müfste. Wir betrachten die symmetrische Funktion

$$\sum_{(i_1, i_2, \ldots i_{n-1})} A_{i_1 i_2 \ldots i_{n-1}} x_1^{i_1} x_2^{i_2} \ldots x_{n-1}^{i_{n-1}}$$

als eine lineare Funktion von x_{n-1}, die ihren Wert nicht ändert, wenn man an x_n Stelle von x_{n-1} setzt. Daraus

folgt dann, dafs der Koeffizient von x_{n-1} verschwindet, also die Funktion in der Form

$$\sum_{(i_1, i_2, \ldots i_{n-2})} A_{i_1 i_2 \ldots i_{n-2} 0} x_1^{i_1} x_2^{i_2} \ldots x_{n-2}^{i_{n-2}}$$

darstellbar ist. Diese sehen wir nun als eine quadratische Funktion der Variablen x_{n-2} an, die ihren Wert nicht ändert, wenn x_{n-1}, x_n an deren Stelle treten, weshalb die Koeffizienten von x_{n-2}^2, x_{n-2} verschwinden müssen. Fährt man so fort, so findet man weiter, dafs die Funktion die Form haben mufs

$$\sum_{(i_1)} A_{i_1 0 \ldots 0} x_1^{i_1}.$$

Da man nun aber $x_2, x_3, \ldots x_n$ an Stelle von x_1 setzen kann, ohne dafs eine Wertänderung erfolgt, so sind alle Koeffizienten mit Ausnahme von $A_{00\ldots0}$ gleich 0, die symmetrische Funktion ist also durch die Grundformen $c_1, c_2, \ldots c_n$ rational darstellbar.

Die Reduktion aller ganzen Funktionen auf die obige Form ist von Nutzen bei der Aufgabe, Funktionen zu ermitteln, die durch gegebene Permutation ihren Wert nicht ändern. Wir wollen dies an einem Beispiele erläutern.

Wir stellen die Aufgabe, alle Funktionen von x_0, x_1, x_2 zu bestimmen, die durch die Permutation

$$S = (x_0, x_1, x_2)$$

nicht geändert werden. Es wird hierbei

$$F(x_0) = x_0^3 - c_1 x_0^2 + c_2 x_0 - c_3$$
$$F_1(x_0, x_1) = x_0^2 + x_0 x_1 + x_1^2 - c_1(x_0 + x_1) + c_2$$
$$F_2(x_0, x_1, x_2) = x_0 + x_1 + x_2 - c_1.$$

Die Funktionen müssen die Form haben

$$\varphi(x_0, x_1, x_2) = a x_0^2 x_1 + b x_0^2 + c x_0 x_1 + d x_0 + e x_1 + f,$$

und es wird dann

$$\varphi_S(x_0, x_1, x_2) = \varphi(x_1, x_2, x_0)$$
$$= a x_1^2 x_2 + b x_1^2 + c x_1 x_2 + d x_1 + e x_2 + f.$$

Nun ergiebt sich aber

$$\begin{aligned}
x_1^2 x_2 &= x_0^2 x_1 - c_1 x_0 x_1 + c_2 x_1 \\
x_1^2 &= -x_0^2 - x_0 x_1 + c_1 x_0 + c_1 x_1 - c_2 \\
x_1 x_2 &= x_0^2 - c_1 x_0 + c_2 \\
x_2 &= -x_0 - x_1 + c_1.
\end{aligned}$$

Setzt man diese Werte in φ_S ein und vergleicht sodann die Koeffizienten mit denen von φ, so erhält man die Gleichungen

$$b = c - b, \quad c = -a c_1 - b, \quad d = b c_1 - c c_1 - e$$
$$e = a c_2 + b c_1 - e + d, \quad 0 = -b c_2 + c c_2 + e c_1,$$

und hieraus folgt

$$b = -\frac{a}{3} c_1, \quad c = -\frac{2a}{3} c_1, \quad d = \frac{a}{3}(c_1^2 - c_2), \quad e = \frac{a}{3} c_2,$$

so dafs φ, wenn man von dem willkürlichen Faktor $\frac{a}{3}$ und einer additiven Konstanten absieht, die folgende Gestalt erhält

$$\varphi(x_0, x_1, x_2) = 3x_0^2 x_1 - c_1(x_0^2 + 2x_0 x_1) + c_1^2 x_0 - c_2(x_0 - x_1).$$

Setzt man nun noch an Stelle von c_1, c_2 die Werte $x_0 + x_1 + x_2, x_0 x_1 + x_1 x_2 + x_2 x_0$, so erhält man

$$\varphi(x_0, x_1, x_2) = x_0^2 x_1 + x_1^2 x_2 + x_2^2 x_0.$$

Verlangt man von der Funktion, dafs sie ihr Vorzeichen ändern soll bei Anwendung der Permutation (x_1, x_2), so erhält man mit leichter Mühe den Ausdruck

$$6x_0^2 x_1 - 2c_1 x_0^2 - 4c_1 x_0 x_1 + 2(c_1^2 - c_2) x_0$$
$$+ 2c_2 x_1 - c_1 c_2 + 3c_3,$$

abgesehen von einer multiplikativen Konstanten.

---

§ 102. Galoisscher Körper einer algebraischen Gleichung.

Jede ganze rationale Funktion der Unbestimmten $x_1, x_2, \dots x_n$ läfst sich in Bezug auf das im vorigen Paragraphen behandelte symmetrische Modulsystem auf eine

lineare Form von $n!$ Gröſsen mit Koeffizienten des Rationalitätsbereiches $(c_1, c_2, \ldots c_n)$ reduzieren. Das Produkt zweier solcher Funktionen läſst sich natürlich ebenfalls auf eine solche Form bringen. Alle rationalen Funktionen von $x_1, x_2, \ldots x_n$ bilden also einen Körper in dem in § 99 definierten Sinne. Es muſs daher nicht nur möglich sein, aus den Gröſsen des Körpers linear unabhängige Gröſsen zu bestimmen, durch die die übrigen homogen und linear ausgedrückt werden können, sondern auch ferner aus ihnen eine primitive Gröſse t abzuleiten, die einer irreduktibeln Gleichung $m$ten Grades $P(t) = 0$ genügt, und durch die alle andern rational ausgedrückt werden können. Eine solche Gröſse t nennt man eine Galoissche Resolvente.

Nehmen wir nun an, diese primitive Gröſse t lasse sich durch $x_1, x_2, \ldots x_n$ in der Form

$$t = \varphi(x_1, x_2, \ldots x_n)$$

ausdrücken, und bilden wir sodann das Produkt

$$\Phi(t) = \prod_{(i_1, i_2, \ldots i_n)} (t - \varphi(x_{i_1}, x_{i_2}, \ldots x_{i_n})),$$

erstreckt über sämtliche Permutationen $i_1, i_2, \ldots i_n$ der n Zahlen $1, 2, \ldots n$, so erhalten wir eine ganze Funktion von t vom Grade $n!$, deren Koeffizienten als symmetrische Funktionen von $x_1, x_2, \ldots x_n$ dem Rationalitätsbereich $(c_1, c_2, \ldots c_n)$ angehören. Diese Gleichung $\Phi(t) = 0$ hat nun aber mit der irreduktibeln Gleichung $P(t) = 0$ eine Wurzel, nämlich

$$t = \varphi(x_1, x_2, \ldots x_n)$$

gemeinschaftlich. Daher muſs $\Phi(t)$ durch $P(t)$ teilbar sein, und die Gleichung $P(t) = 0$ hat also nur Wurzeln von der Form

$$\varphi(x_{k_1}, x_{k_2}, \ldots x_{k_n}),$$

wo $k_1, k_2, \ldots k_n$ durch gewisse Permutationen aus den Zahlen $1, 2, \ldots n$ hervorgehen. Da nun aber $x_1, x_2, \ldots x_n$ alle rational durch t darstellbar sind, so sind es auch die Wurzeln $\varphi(x_{k_1}, x_{k_2}, \ldots x_{k_n})$ der Gleichung $P(t) = 0$. Wir erhalten somit folgenden Fundamentalsatz:

Alle Wurzeln der Gleichung $P(t) = 0$ gehen durch rationale Transformationen aus einer einzigen hervor.

Beachten wir ferner, dafs die Wurzeln $x_1, x_2, \ldots x_n$ der Gleichung $F(x) = 0$ rational durch t bestimmt werden können in der Form

$$x_1 = \chi_1(t), \quad x_2 = \chi_2(t), \ldots x_n = \chi_n(t),$$

und dafs $F(x)$, wenn der Koeffizient der höchsten Potenz der Einheit gleich ist, zerlegt werden kann in das Produkt

$$F(x) = (x - x_1)(x - x_2) \ldots (x - x_n)$$
$$= (x - \chi_1(t))(x - \chi_2(t)) \ldots (x - \chi_n(t)),$$

so hat die Differenz $F(x) - \prod_{i=1}^{i=n}(x - \chi_i(t))$ mit $P(t) = 0$ eine Wurzel gemeinsam und mufs daher durch $P(t)$ teilbar sein. Wir können also dem Fundamentalsatz der Algebra (§ 97) auch eine von der Einführung irrationaler Gröfsen völlig unabhängige Fassung geben:

Man kann jede ganze Funktion $F(x)$ in Bezug auf $P(t)$ als Modul in ein Produkt von lauter Linearfaktoren zerlegen, die rational von t abhängig sind.

---

§ 103. Gruppeneigenschaft der rationalen Transformationen eines Körpers.

Die Darlegungen im vorigen Paragraphen haben uns zu irreduktibeln Gleichungen geführt, deren sämtliche Wurzeln rational durch eine einzige darstellbar sind. Solche Gleichungen wollen wir jetzt in allgemeinerer Form betrachten.

Wird an die Stelle des Argumentes x der Primfunktion $P(x)$ eine rationale Funktion $\varphi(x)$ gesetzt, so ist $P\varphi(x)$ entweder durch $P(x)$ teilbar oder teilerfremd zu $P(x)$. Tritt das erstere ein, so nennen wir $\varphi(x)$ kurz eine rationale Transformation von $P(x)$. Eine solche Transformation giebt es stets, nämlich $\varphi(x) = x$, die man auch die identische Transformation nennt (§ 14). Aus einer einzigen lassen sich aber unendlich viele ableiten. Ist nämlich

$$\varphi(x) \equiv \psi(x) \mod P(x),$$

so ist auch stets allgemein

$$F\varphi(x) \equiv F\psi(x) \mod P(x),$$

wo F eine beliebige Funktion bedeutet, für die wir dann P setzen können, und man kann offenbar zu $\varphi(x)$ unendlich viele nach $P(x)$ kongruente Funktionen $\psi(x)$ herstellen, indem man zu $\varphi(x)$ ein beliebiges Vielfaches von $P(x)$ hinzufügt. Es hat aber kein Interesse, diese alle mit in Betracht zu ziehen, und wir werden uns daher auf inkongruente Transformationen beschränken. Da jede gebrochene Funktion nach dem Modul $P(x)$ auch durch eine ihr kongruente ganze Funktion ersetzt werden kann, so kann man die sämtlichen inkongruenten Transformationen, wenn man will, als ganz voraussetzen.

Einige einfache Beispiele mögen zunächst das Auftreten von rationalen Transformationen erläutern.

Ist $P(x)$ eine quadratische Funktion

$$P(x) = x^2 + a x + b,$$

so giebt es aufser der identischen Transformation noch die Transformation $\varphi(x) = -x - a$, da

$$P\varphi(x) = \varphi(x)^2 + a\varphi(x) + b = (x + a)^2 - a(x + a) + b$$
$$= x^2 + a x + b = P(x)$$

wird. Sind die Koeffizienten einer Primfunktion

$$P(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n$$

so beschaffen, dafs $a_0 = a_n$, $a_1 = a_{n-1}$ u. s. w., allgemein

$$a_i = a_{n-i} \qquad (i = 0, 1, \ldots n)$$

ist, so ist aufser der identischen $\varphi(x) = \frac{1}{x}$ eine zu $P(x)$ gehörige Transformation, denn es wird

$$x^n P\varphi(x) = a_0 + a_1 x + \ldots + a_{n-1} x^{n-1} + a_n x^n = P(x),$$

und da x^n nicht durch $P(x)$ teilbar ist, so mufs

$$P\varphi(x) \equiv 0 \text{ mod } P(x)$$

sein. Hat $P(x)$ die angegebene Form, so nennt man die Gleichung $P(x) = 0$ auch eine reciproke Gleichung.

Ist p eine Primzahl, so ist die Funktion

$$P(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \ldots + x + 1$$

irreduktibel, wie in § 60 festgestellt wurde. Ist nun n teilerfremd zu p, so stellt $\varphi(x) = x^n$ eine rationale Transformation von P(x) dar, denn es ist

$$x^{np} - 1 = (x^n - 1)(x^{n(p-1)} + \ldots + x^n + 1) = (x^n - 1)P(x^n)$$

durch $x^p - 1$ und somit auch durch P(x) teilbar; da aber $(x^n - 1, x^p - 1) = x - 1$ (§ 63) und daher $x^n - 1$ zu P(x) teilerfremd ist, so mufs

$$P(x^n) \equiv 0 \mod P(x)$$

sein. Giebt man n der Reihe nach die Werte $1, 2, \ldots p - 1$, so erhält man $(p - 1)$ Transformationen

$$x, x^2, \ldots x^{p-2}, x^{p-1},$$

die auch inkongruent nach dem Modul P(x) sind, weil aus

$$x^a \equiv x^b \mod P(x)$$

notwendig

$$a \equiv b \mod p$$

folgt und umgekehrt. Nämlich wenn $a > b$ ist, so würde $x^b(x^{a-b} - 1)$ und also auch $x^{a-b} - 1$ durch P(x) teilbar sein. Da aber $x^{a-b} - 1$ den Teiler $x - 1$ besitzt, den P(x) nicht hat, so müfste auch $x^{a-b} - 1$ durch $x^p - 1$ teilbar sein, was nur geschehen kann, wenn $a - b$ durch p teilbar ist. Übrigens ist noch leicht zu erkennen, dafs die Gleichung $P(x) = \frac{x^p - 1}{x - 1} = 0$ eine reciproke, und dafs $x^{p-1} \equiv \frac{1}{x} \mod P(x)$ ist.

Wenden wir uns jetzt wieder dem allgemeinen Falle zu, und ziehen wir die sämtlichen rationalen Transformationen von P(x) in Betracht, so ist es nicht schwierig, die folgende wichtige Eigenschaft festzustellen:

Sind $\varphi(x)$ und $\psi(x)$ zwei Transformationen von P(x), so sind auch $\varphi\psi(x)$ und $\psi\varphi(x)$ solche.

Mit $\varphi\psi$ und $\psi\varphi$ bezeichnen wir die aus der Zusammensetzung von φ und ψ hervorgehenden Funktionen (§ 14). Der Beweis ist sehr einfach. Aus den Kongruenzen

$$P\varphi(x) \equiv 0, \quad P\psi(x) \equiv 0 \mod P(x)$$

ergiebt sich zunächst, wenn man in der ersteren das Argument x durch $\psi(x)$, in der zweiten durch $\varphi(x)$ ersetzt,

$$P\varphi\psi(x) \equiv 0 \mod P\psi(x), \quad P\psi\varphi(x) \equiv 0 \mod P\varphi(x),$$

und hieraus folgt in Verbindung mit den soeben hingeschriebenen Kongruenzen, dafs erst recht

$$P\varphi\psi(x) \equiv 0, \quad P\psi\varphi(x) \equiv 0 \mod P(x)$$

ist, und damit ist die Richtigkeit des Satzes dargethan. Mit Hilfe dieses Satzes sind wir imstande, aus zwei gegebenen Transformationen oder vielmehr schon aus einer einzigen durch wiederholte Zusammensetzung neue zu erzeugen. Es läfst sich nun aber zeigen, dafs es nur eine beschränkte Anzahl von inkongruenten Transformationen von $P(x)$ giebt. Ist nämlich $\varphi(x)$ eine Transformation von $P(x)$, so ist, da $P(t) - P\varphi(x)$ durch $t - \varphi(x)$ teilbar ist, $P(t)$ durch $t - \varphi(x)$ teilbar nach dem Modul $P(x)$. Sind nun mehrere inkongruente Transformationen

$$\varphi_0(x) = x, \ \varphi_1(x), \ldots \ \varphi_{m-1}(x)$$

gegeben, so sind von den Linearfaktoren

$$t - x, \ t - \varphi_1(x), \ldots \ t - \varphi_{m-1}(x)$$

immer je zwei teilerfremd zu einander, und daraus folgt dann leicht, dafs $P(t)$ abgesehen von Vielfachen von $P(x)$ durch das Produkt

$$(t - x)(t - \varphi_1(x)) \ldots (t - \varphi_{m-1}(x))$$

teilbar ist. Es kann aber die Anzahl der Faktoren und daher auch die der inkongruenten Transformationen niemals den Grad von $P(x)$ überschreiten. Später werden wir noch zu einer genaueren Angabe über die Anzahl der Transformationen gelangen. Vorläufig genügt es uns schon, ihre Endlichkeit festgestellt zu haben.

Sind nun von $P(x)$ eine Reihe von inkongruenten Transformationen gegeben, so kann man aus diesen durch wiederholte Zusammensetzung immer neue ableiten, bis man schliefslich zu einem Abschlufs eines Systems gelangt, das nicht mehr vervollständigt werden kann. Ein solches System nennen wir dann eine Gruppe von Transformationen. Um diese weiter zu untersuchen, machen wir die folgende Bemerkung.

Sind φ, ψ und χ drei Transformationen von P(x), und ist

$$\chi\varphi(x) \equiv \psi\varphi(x) \mod P(x),$$

so mufs auch

$$\chi(x) \equiv \psi(x) \mod P(x)$$

sein. Denn wäre dies nicht der Fall, so würde, da P(x) eine Primfunktion ist, die Differenz $\chi(x) - \psi(x)$ zu P(x) teilerfremd sein müssen. Ersetzen wir x durch $\varphi(x)$, so ergäbe sich hieraus, dafs dann auch $\chi\varphi(x) - \psi\varphi(x)$ zu $P\varphi(x)$ teilerfremd wäre. Das ist aber sicher nicht der Fall, da die Differenz dieser beiden Funktionen der Annahme gemäfs durch P(x) teilbar ist.

Ist nun eine Gruppe von inkongruenten Transformationen

$$x, \varphi_1(x), \varphi_2(x), \ldots \varphi_{m-1}(x)$$

gegeben, und setzt man alle diese Transformationen mit einer φ von ihnen zusammen, so erhält man das System

$$\varphi(x), \varphi_1\varphi(x), \ldots \varphi_{m-1}\varphi(x).$$

Dieses stimmt, von der Reihenfolge der Transformationen abgesehen, vollständig mit dem ursprünglichen System überein; denn erstens kann es nur, weil das System vollständig war, Transformationen enthalten, die in ihm schon vertreten waren, und zweitens sind zufolge des vorhergehenden Satzes auch alle Transformationen inkongruent. Daher giebt es in der Gruppe eine Transformation $\bar{\varphi}(x)$, so dafs

$$\bar{\varphi}\varphi(x) \equiv x \mod P(x)$$

wird. Diese Transformation nennt man die inverse Transformation zu $\varphi(x)$. Aus der Kongruenz folgt

$$\varphi(\bar{\varphi}\varphi(x)) \equiv (\varphi\bar{\varphi})\varphi(x) \equiv \varphi(x) \mod P(x)$$

und daraus nach dem schon einmal benutzten Satze

$$\varphi\bar{\varphi}(x) \equiv x \mod P(x),$$

so dafs also $\varphi(x)$ umgekehrt die inverse Transformation von $\bar{\varphi}(x)$ ist.

Man kann aber, und zwar durch ähnliche Schlufsweise, das erhaltene Resultat noch verallgemeinern zu dem Satze:

Sind φ, ψ zwei Transformationen von P(x), so giebt es immer eine dritte χ in der Gruppe, so dafs

$$\varphi(x) \equiv \psi\chi(x) \bmod P(x)$$

ist.

Aber es ist leichter, diesen Satz daraus abzuleiten, dafs die diese Bedingung erfüllende Transformation $\chi(x)$ sich einfach in der Form $\bar{\psi}\varphi(x)$ darstellen läfst, die in der Gruppe enthalten ist, weil φ, ψ und daher auch $\bar{\psi}$ es sind.

Schon aus einer einzigen Transformation $\varphi(x)$ kann man durch wiederholte Zusammensetzung eine Gruppe erzeugen. Ist n die kleinste Zahl, für die

$$\varphi^n(x) \equiv x \bmod P(x)$$

wird, so ist die Gruppe von der Ordnung n, denn sie besitzt die n inkongruenten Transformationen

$$x, \varphi(x), \varphi^2(x), \ldots \varphi^{n-1}(x).$$

Die Zahl n nennt man auch die Ordnung der Transformation $\varphi(x)$, und man hat allgemein

$$\varphi^a\varphi^b(x) \equiv \varphi^c(x) \bmod P(x),$$

sobald

$$a + b \equiv c \bmod n$$

ist.

Wenn eine Gruppe G alle Transformationen einer Gruppe H enthält, die durch

$$\varphi_0(x) = x, \ \varphi_1(x), \ldots \varphi_{r-1}(x)$$

bezeichnet werden mögen, so lassen sich aus G eine Reihe von Transformationen $\psi_1(x), \psi_2(x), \ldots \psi_{n-1}(x)$ bestimmen, so dafs durch die Transformationen von H im Verein mit den folgenden

$$\begin{array}{lll} \psi_1(x), & \varphi_1\psi_1(x), & \ldots \varphi_{r-1}\psi_1(x) \\ \psi_2(x), & \varphi_1\psi_2(x), & \ldots \varphi_{r-1}\psi_2(x) \\ \vdots & & \\ \psi_{n-1}(x), & \varphi_1\psi_{n-1}(x), & \ldots \varphi_{r-1}\psi_{n-1}(x) \end{array}$$

alle Transformationen von G, jede nur einmal, zur Darstellung gelangen. Den Beweis hierfür brauchen wir hier nicht mehr zu erbringen, da in § 26 und § 37 schon ähnliche Schlüsse gebraucht wurden. Es ergiebt sich daraus der Satz:

Enthält eine Gruppe G alle Transformationen einer Gruppe H, so ist die Ordnung von G ein Vielfaches der Ordnung von H.

Den Quotienten $\frac{m}{r} = n$ der beiden Ordnungen nennt man auch den Index der Gruppe H in Bezug auf G.

Betrachten wir alle Transformationen von G, die in dem obigen Schema in derselben Zeile stehen, also

$$\psi_i(x),\ \varphi_1 \psi_i(x), \ldots \varphi_{r-1} \psi_i(x), \quad (i = 0, 1, \ldots n-1)$$

so ersehen wir, dafs sie alle aus einer nicht in H enthaltenen Transformation ψ_i durch Zusammensetzung mit den Transformationen der Gruppe H hervorgehen; man kann daher auch zwei beliebige aus einander durch Zusammensetzung mit einer Transformation von H ableiten, denn es ist

$$\varphi_k \psi_i = (\varphi_k \bar{\varphi}_h) \varphi_h \psi_i.$$

Eine solche Herleitung ist nicht möglich bei Transformationen, die verschiedenen Zeilen angehören, denn wäre z. B.

$$\varphi' \psi_i = \varphi'' \psi_k, \qquad (i \neq k)$$

so müfste

$$\psi_i = \bar{\varphi}' \varphi'' \psi_k$$

sein, also ψ_i in derselben Zeile wie ψ_k vorkommen. Nennen wir zwei Transformationen von G, die sich durch Zusammensetzung mit einer Transformation aus H herleiten lassen, äquivalent in Bezug auf H, die, bei denen das nicht der Fall ist, konjugiert, so bilden die Transformationen

$$x,\ \psi_1(x), \ldots \psi_{n-1}(x)$$

ein vollständiges System konjugierter Transformationen von G in Bezug auf H. Da $\psi_i \psi_k$ in G enthalten, also in der Form $\varphi_g \psi_h$ darstellbar ist, so ist dann $\psi_i \psi_k$ äquivalent ψ_h.

---

§ 104. Rationale Transformationen von Funktionen eines Körpers.

Es sei wieder P(x) eine Primfunktion und eine Gruppe G von rationalen Transformationen von P(x) gegeben. F(x) sei eine beliebige Funktion, und wir stellen uns die Aufgabe, aus der Gruppe der rationalen Transformationen $\varphi(x)$ von P(x) diejenigen auszusondern, die F(x) nach dem Modul P(x) ungeändert lassen, also der Kongruenz

$$F\varphi(x) \equiv F(x) \mod P(x)$$

genügen. Wir nennen sie die zu F(x) gehörigen Transformationen von P(x). Es läfst sich sofort zeigen, dafs diese eine Untergruppe von G bilden. Sind nämlich $\varphi(x)$ und $\psi(x)$ zwei Transformationen von den gesuchten Eigenschaften, ist also

$$F\varphi(x) \equiv F(x),\ F\psi(x) \equiv F(x) \mod P(x),$$

so folgt

$$F\varphi\psi(x) \equiv F\psi(x) \mod P\psi(x)$$
$$F\psi\varphi(x) \equiv F\varphi(x) \mod P\varphi(x).$$

Da aber $\varphi(x)$ und $\psi(x)$ Transformationen von P(x) sind, so ist

$$P\varphi(x) \equiv 0,\ P\psi(x) \equiv 0 \mod P(x),$$

und so ergiebt sich, dafs

$$F\varphi\psi(x) \equiv F\psi\varphi(x \equiv F\psi(x) \equiv F\varphi(x) \equiv F(x) \mod P(x)$$

ist. Also gehören $\varphi\psi$ und $\psi\varphi$ ebenfalls zu F(x) und wir erhalten den Satz:

Alle zu einer beliebigen Funktion F(x) gehörigen Transformationen von P(x) bilden eine in G enthaltene Gruppe H.

Es ist ferner klar, dafs zu jeder Gruppe H von rationalen Transformationen von P(x) auch Funktionen existieren, die durch sie ihren Wert nicht ändern. Man braucht ja nur irgend welche symmetrische Funktionen aller rationalen Transformationen der Gruppe H zu bilden, um solche zu erhalten. Hat man zwei solche Funktionen gebildet, so hat man in ihrer Summe und ihrem Produkt

ebenfalls Funktionen von derselben Eigenschaft. Hieraus folgt aber:

Alle Funktionen von x, die durch eine Gruppe H rationaler Transformationen von P(x) sich nach dem Modul P(x) nicht ändern, bilden einen Körper. Wir nennen ihn den zur Gruppe H gehörigen Körper.

Ist der Grad von P(x) gleich m und die Ordnung der Gruppe H gleich r, so kann man den Grad n dieses Körpers auf folgende Weise ermitteln. Wir nehmen an, dafs $\theta(x)$ ein primitives Element dieses Körpers sei, und dafs x der irreduktibeln Gleichung in t

$$\Phi(t; \theta(x)) = 0$$

genügt, oder dafs

$$\Phi(x; \theta(x)) \equiv 0 \mod P(x)$$

ist. Ist dann $\varphi_i(x)$ eine Transformation der Gruppe, so folgt, da $\theta\varphi_i(x) \equiv \theta(x)$ ist, dafs auch

$$\Phi(\varphi_i(x); \theta(x)) \equiv 0 \mod P(x) \quad (i=0,1,\ldots r-1)$$

ist. Nun ergiebt sich sofort, dafs $\Phi(t)$ nach dem Modul P(x) durch das Produkt

$$\prod_{i=0}^{i=r-1}(t-\varphi_i(x))$$

teilbar ist, das über sämtliche Transformationen der Gruppe erstreckt wird, denn die sämtlichen Transformationen sind inkongruent, die entsprechenden Linearfaktoren $t-\varphi_i(x)$ ebenfalls und daher teilerfremd. Der Grad der Funktion $\Phi(t; \theta(x))$ ist also mindestens gleich r in Bezug auf t. Er kann aber auch nicht gröfser sein, da das Produkt

$$\prod_{i=0}^{i=r-1}(t-\varphi_i(x))$$

schon eine rationale Funktion von t darstellt, deren Koeffizienten als symmetrische Funktionen der sämtlichen Transformationen schon in dem zur Gruppe H gehörigen Körper enthalten sind. Alle Gröfsen des durch die Gleichung $P(x) = 0$ definierten Körpers lassen sich daher durch das Fundamentalsystem

$$x^{i-1}\,\theta(x)^{k-1} \quad (i = 1, 2, \ldots r; k = 1, 2, \ldots n)$$

darstellen. Daher ist

$$m = r\,n, \quad n = \frac{m}{r}.$$

Die Ordnung r jeder Gruppe H ist also ein Teiler des Grades m des durch $P(x) = 0$ definierten Körpers, und der komplementäre Teiler n stellt den Grad des zur Gruppe H gehörigen Teilkörpers dar.

Beispiel. Wenden wir dies auf die reciproken Gleichungen an, die wir im vorigen Paragraphen besprochen haben, so ergiebt sich zunächst, weil die Transformation $\varphi(x) = \frac{1}{x}$ die Ordnung 2 hat, dafs der Grad m einer irreduktibeln reciproken Gleichung gerade ist. Nehmen wir $m = 2n$ an, und beachten wir, dafs wir als Fundamentalsystem auch die $2n$ Potenzen

$$x^i, \; \frac{1}{x^i} \qquad (i = 1, 2, \ldots n)$$

zu Grunde legen können, so läfst sich in den n Gröfsen

$$y_i = x^i + \frac{1}{x^i} \qquad (i = 1, 2, \ldots n)$$

ein Fundamentalsystem des zur Gruppe gehörigen Körpers erkennen. Denn damit eine Gröfse des Gesamtkörpers, die sich immer in der Form

$$f(x) = \sum_i (u_i\, x^i + v_i \frac{1}{x^i}) \qquad (i = 1, 2, \ldots n)$$

läfst, durch die Transformation $\varphi(x) = \frac{1}{x}$ nicht geändert wird, ergiebt sich als notwendige Bedingung $u_i = v_i$. Man beweist ferner leicht, dafs die Gröfse

$$y = x + \frac{1}{x}$$

eine primitive Gröfse des Körpers ist. Aus den Rekursionsformeln

$$y_{i+1} = y_i\, y - y_{i-1}$$

folgt, dafs sich y_i linear und homogen durch $1, y, \ldots y^i$ darstellen läfst, während andererseits aus

$$y^i = (x + \frac{1}{x})^i = y_i + i_1 y_{i-1} + i_2 y_{i-2} + \dots$$

(wo $i_1, i_2, \dots$ Binomialkoeffizienten bedeuten) hervorgeht, dafs y^i sich ebenso durch $y_1, y_2, \dots y_i$ und die Einheit ausdrücken läfst. Die Gröfse y genügt also einer irreduktiblen Gleichung $n$ten Grades. Die Auflösung der reciproken Gleichungen vom Grade $m = 2n$ kann daher auf die Lösung einer Gleichung $n$ten Grades zurückgeführt werden. Ist diese gelöst, so ist x durch eine quadratische Gleichung

$$x^2 - yx + 1 = 0$$

zu ermitteln.

---

§ 105. Zerlegung Galoisscher Funktionen.

Eine besonders einfache Gestalt gewinnt der Hauptsatz des vorigen Paragraphen, wenn man annimmt, dafs $P(x)$ eine Galoissche Funktion ist, also die Ordnung der Gruppe G ihrer rationalen Transformationen mit dem Grade m von $P(x)$ übereinstimmt. Bilden dann die Transformationen

$$x, \varphi_1(x), \dots \varphi_{r-1}(x)$$

eine in G enthaltene Untergruppe H, so ist der Grad des zu H gehörigen Körpers gleich dem Index der Gruppe H in Bezug auf G.

Eine primitive Gröfse dieses zu H gehörigen Körpers läfst sich dann durch den folgenden Satz charakterisieren:

Die hinreichende und notwendige Bedingung dafür, dafs eine Gröfse $\theta(x)$ eine primitive Gröfse des zur Gruppe H gehörigen Körpers ist, besteht darin, dafs die konjugierten Funktionen

$$\theta(x), \theta\psi_1(x), \dots \theta\psi_{n-1}(x)$$

verschieden (nach dem Modul $P(x)$) sind, wenn

$$x, \psi_1(x), \psi_2(x), \dots \psi_{n-1}(x)$$

ein vollständiges System konjugierter Transformationen von G in Bezug auf H bedeutet.

Die Bedingung ist notwendig, weil sonst $\theta(x)$ noch durch andere Transformationen als diejenigen von H unverändert bliebe, also zu einer die Gruppe H enthaltenden Gruppe H' gehören würde, deren Index n' kleiner als n wäre; dann aber wäre $\theta(x)$ eine algebraische Gröfse von höchstens n' ter Ordnung. Andrerseits ist aber auch diese Bedingung hinreichend. Denn ist

$$Q(y) = 0$$

die irreduktible Gleichung für $\theta(x)$, so folgt aus

$$Q\,\theta(x) \equiv 0 \mod P(x)$$

sofort

$$Q\,\theta\,\psi_i(x) \equiv 0 \mod P(x), \quad (i = 1, 2, \dots n-1)$$

und da die Gröfsen

$$y - \theta(x),\ y - \theta\,\psi_1(x), \dots y - \theta\,\psi_{n-1}(x)$$

nach Annahme alle inkongruent mod P(x) sind, so mufs Q(y) durch das Produkt

$$\Pi(y - \theta\,\psi_i(x)) \mod P(x) \quad (i = 0, 1, \dots n-1)$$

teilbar sein, das eine ganze Funktion m ten Grades von y mit rationalen Koeffizienten darstellt.

Wird die zur Gruppe H gehörige primitive Gröfse $\theta(x)$ dem Körper adjungiert, so wird die Galoissche Funktion P(x) irreduktibel, der irreduktible Faktor $P_1(x, \theta(x))$ besitzt nur noch die Transformationen der Gruppe H und ist wieder eine Galoissche Funktion.

So verhält sich die Galoissche Funktion P(x), wenn man irgend eine rationale Funktion von x adjungiert. Wir müssen jetzt noch untersuchen, wie sie sich verhält, wenn irgend eine algebraische Gröfse adjungiert wird. Das ist aber sehr einfach zu beurteilen, da sofort ersichtlich ist, dafs jeder irreduktible Faktor dann wieder eine Galoissche Funktion darstellen mufs. Denn da P(x) soviele rationale Transformationen besitzt, wie der Grad angiebt, so müssen sich diese auf die irreduktibeln Faktoren ihrem Grade nach verteilen, wobei die Gruppeneigenschaft wegen der Irreduktibilität von Bestand bleiben mufs. Zerfällt also P(x) durch Adjunktion einer durch die Gleichung Q(y) = 0 definierten algebraischen Gröfse y, und ist $P_1(x, y)$

ein irreduktibler Faktor, so mufs dieser eine Gruppe H von Transformationen besitzen

$$x, \varphi_1(x), \dots \varphi_{r-1}(x),$$

also für jeden Wert von t, abgesehen von einer etwa auftretenden multiplikativen Konstanten,

$$P_1(t, y) = (t - x)(t - \varphi_1(x)) \dots (t - \varphi_{r-1}(x))$$

sein. Nun ist aber leicht nachweisbar, dafs die auf der rechten Seite stehende Gröfse bei geeigneter Wahl von t eine primitive Gröfse $\Theta(x)$ des zu H gehörigen Körpers darstellt. Denn von den durch ein System konjugierter Transformationen von G in Bezug auf H

$$x, \psi_1(x), \dots \psi_{n-1}(x)$$

hervorgehenden konjugierten Funktionen

$$\Theta(x), \Theta\psi_1(x), \dots \Theta\psi_{n-1}(x)$$

sind sicher nicht je zwei für jeden Wert von t gleich, da $\Theta\psi_i(x)$ für $t = \psi_i(x)$ verschwindet, während $\Theta\psi_k(x) = \prod_h (\psi_k(x) - \varphi_h\psi_i(x))$ es nicht thut, wenn $i \neq k$ ist. Man kann daher t als rationale Zahl auf unendlich mannigfaltige Art (§ 62) so bestimmen, dafs alle konjugierten Funktionen verschieden sind, und unter dieser Voraussetzung ist dann $\Theta(x)$ ein primitives Element des zu H gehörigen Körpers. Dann aber folgt nicht nur, dafs die Zerfällung von P(x) sich durch Adjunktion einer Gröfse $\Theta(x)$ des Körpers erreichen läfst, sondern aus der Gleichung

$$P_1(t, y) = \Theta(x)$$

ergiebt sich weiterhin, dafs $\Theta(x)$ rational durch y dargestellt werden kann, also der durch die Gleichung $Q(y) = 0$ definierte Körper den zur Gruppe H gehörigen enthalten mufs.

Ist der Grad von Q(y) eine Primzahl, so folgt hieraus noch, dafs die beiden Körper identisch sein müssen.

Ist $\Theta(x)$ eine primitive Gröfse des zur Gruppe H gehörigen Körpers, so sind, wie wir oben gesehen haben, die den sämtlichen konjugierten Transformationen von G in Bezug auf H entsprechenden konjugierten Gröfsen

$$\Theta(x), \Theta\psi_1(x), \dots \Theta\psi_{n-1}(x)$$

alle von einander verschieden. Zu ihnen gehören ebenfalls gewisse Gruppen

$$H, H_1, H_2, \ldots H_{n-1}$$

von Transformationen, die man als konjugierte Gruppen bezeichnet. Ihre Bestimmung kann auf ähnlichem Wege erfolgen wie in § 38. Ist nämlich für irgend eine in G enthaltene Transformation $\varrho(x)$

$$\theta\, \psi_i\, \varrho\,(x) = \theta\, \psi_i\,(x), \qquad (i = 0, 1, \ldots n-1)$$

so folgt durch Einsetzung von $\overline{\psi}_i(x)$ auf beiden Seiten

$$\theta\, \psi_i\, \varrho\, \overline{\psi}_i\,(x) = \theta\,(x),$$

dafs also $\psi_i\, \varrho\, \overline{\psi}_i\,(x)$ als eine zu $\theta\,(x)$ gehörige Transformation in H enthalten ist, und wenn demgemäfs

$$\psi_i\, \varrho\, \overline{\psi}_i\,(x) = \varphi_h\,(x)$$

gesetzt wird, so erhält man

$$\varrho\,(x) = \overline{\psi}_i\, \varphi_h\, \psi_i\,(x)$$

als allgemeine Form der zu $\theta\, \psi_i\,(x)$ gehörigen Transformationen. Lassen wir h die Werte $0, 1, \ldots r-1$ annehmen, so bilden die sich ergebenden Transformationen

$$x,\ \overline{\psi}_i\, \varphi_1\, \psi_i,\ \overline{\psi}_i\, \varphi_2\, \psi_i, \ldots \overline{\psi}_i\, \varphi_{r-1}\, \psi_i \qquad (i = 0, 1, \ldots n-1)$$

in der That eine Gruppe H_i, da durch Zusammensetzung je zweier immer wieder eine in der Reihe enthaltene Transformation hervorgeht (vgl. § 38).

Wenn alle diese Gruppen $H, H_1, H_2, \ldots H_{n-1}$ mit einander übereinstimmen, so nennt man H eine ausgezeichnete Gruppe von G. Da die konjugierten Funktionen dann alle primitive Gröfsen in H sind, so lassen sie sich alle rational durch eine einzige von ihnen darstellen, und der zu ihnen gehörige durch die Gleichung $Q\,(y) = 0$ definierte Körper ist daher wieder ein Galoisscher. Aber auch das Umgekehrte gilt. Die Transformationen wollen wir jetzt genauer untersuchen und setzen

$$\theta\, \psi_i\,(x) = \Psi_i\, \theta\,(x). \qquad (i = 1, 2, \ldots n-1)$$

Führen wir auf beiden Seiten die Transformation $\psi_k(x)$ aus, so folgt

$$\Theta\,\psi_i\,\psi_k(x) = \Psi_i\,\Theta\,\psi_k(x) = \Psi_i\,\Psi_k\,\Theta(x).$$

Nun ist $\psi_i\,\psi_k$ in der Gruppe G enthalten, also in der Form $\varphi_g\,\psi_h$ darstellbar und in Bezug auf H der Transformation ψ_h äquivalent. Dann ist andrerseits

$$\Theta\,\psi_i\,\psi_k(x) = \Theta\,\varphi_g\,\psi_h(x) = \Theta\,\psi_h(x) = \Psi_h\,\Theta(x).$$

Wenn also $\psi_i\,\psi_k$ in Bezug auf H äquivalent ψ_h ist, so ist

$$\Psi_i\,\Psi_k\,\Theta(x) = \Psi_h\,\Theta(x),$$

und auch das Umgekehrte läfst sich leicht zeigen. Es bilden also die Funktionen

$$y,\ \Psi_1(y),\ \Psi_2(y), \ldots \Psi_{n-1}(y),$$

eine Gruppe des zu H gehörigen Galoisschen Körpers $Q(y) = 0$, deren Eigenschaften sich aus den konjugierten Transformationen

$$x,\ \psi_1(x),\ \psi_2(x), \ldots \psi_{n-1}(x)$$

von G in Bezug auf H ableiten lassen.

---

§ 106. Abelsche und cyklische Körper.

Wenn alle rationalen Transformationen der Gruppe G des zu $P(x) = 0$ gehörigen Körpers vertauschbar sind, d. h. wenn für irgend zwei Transformationen φ und ψ aus G stets

$$\varphi\,\psi = \psi\,\varphi$$

ist, so ist jede in G enthaltene Gruppe H eine ausgezeichnete; denn weil $\overline{\psi_i}\,\varphi_k\,\psi_i = \varphi_k$ wird, so sind alle konjugierten Gruppen $H, H_1, \ldots H_{n-1}$ dieselben. Eine Galoissche Gleichung mit lauter vertauschbaren Transformationen nennt man auch eine Abelsche Gleichung. Jeder im Abelschen Körper enthaltene Teilkörper $Q(y) = 0$ ist aber nicht nur, wie aus dem soeben Bemerkten hervorgeht, ein Galoisscher Körper, sondern wieder ein Abelscher Körper. Denn da $\psi_i\,\psi_k = \psi_k\,\psi_i$ ist, so folgt für die Transformationen von $Q(y)$ leicht aus der Schlufsbemerkung des vorigen Paragraphen, dafs auch $\Psi_i\,\Psi_k = \Psi_k\,\Psi_i$ wird.

Ein besonders einfacher Fall der Abelschen Körper sind die cyklischen, die man auch nach Gaufs benennen könnte, da dieser ihre Eigenschaften im Anschlufs an die Kreisteilungsgleichungen entwickelt hat, worauf sich dann die Verallgemeinerungen zunächst von Abel und darauf von Galois aufbauten. Ein cyklischer Körper ist dadurch definiert, dafs sich alle seine Transformationen durch Iteration einer einzigen $\varphi(x)$ erzeugen lassen, so dafs die die Gruppe G dargestellt werden kann durch

$$x, \varphi(x), \varphi^2(x), \dots \varphi^{m-1}(x),$$

wobei $\varphi^m(x) \equiv x$, mod $P(x)$ wird. Da dann

$$\varphi^a \varphi^b(x) \equiv \varphi^c(x) \mod P(x)$$

wird, sobald

$$a + b \equiv c \mod m$$

ist, so folgt, dafs die Transformationen

$$\varphi^{a_0}(x), \varphi^{a_1}(x), \dots \varphi^{a_{r-1}}(x)$$

dann und nur dann eine in G enthaltene Gruppe H bilden, wenn die Exponenten

$$a_0, a_1, \dots a_{r-1}$$

eine Gruppe von Resten im vollständigen Restsystem nach dem Modul m bilden. Aus § 25 ergiebt sich aber, dafs diese Gruppe dargestellt werden kann durch

$$0, n, 2n, \dots (r-1)n,$$

wobei r und n zwei beliebige komplementäre Teiler von m sein können. Setzen wir daher

$$\varphi^n(x) = \psi(x),$$

so besteht H aus den Transformationen

$$x, \psi(x), \psi^2(x), \dots \psi^{r-1}(x),$$

und es folgt:

Jede cyklische Gruppe vom Grade m enthält nur cyklische Untergruppen und von jeder Ordnung, die ein Teiler von m ist.

Beachten wir nun, dafs die zu H konjugierten Transformationen sich darstellen lassen durch

$$x, \varphi(x), \varphi^2(x), \dots \varphi^{n-1}(x),$$

so ergiebt sich für die Transformationen des zu H gehörigen Körpers $Q(y) = 0$, wenn $\Theta(x)$ eine primitive Gröfse in diesem und

$$\Theta\varphi(x) = \Phi\,\Theta(x)$$

ist, dafs dessen sämtliche Transformationen

$$y, \Phi(y), \Phi^2(y), \ldots \Phi^{n-1}(y)$$

ebenfalls einen cyklischen Körper bedingen, oder:

Jeder in einem cyklischen Körper enthaltene Teilkörper ist wieder cyklisch.

Wir wollen jetzt die Frage beantworten: Wann zerfällt eine Galoissche Gleichung $P(x) = 0$ durch Adjunktion cyklischer Körper $Q(y) = 0$? Wir können dabei annehmen, dafs dieser cyklische Körper im Galoisschen enthalten ist, denn wir haben bewiesen, dafs jede Zerfällung auch durch Adjunktion einer Gröfse des Galoisschen Körpers herbeigeführt werden kann.

Ist also $\Theta(x)$ eine primitive Gröfse eines solchen cyklischen Körpers, H die Gruppe der $\Theta(x)$ gehörigen Transformationen

$$x, \varphi_1(x), \varphi_2(x), \ldots \varphi_{r-1}(x),$$

so mufs diese Gruppe in G als ausgezeichnete enthalten sein. Die konjugierten Funktionen zu $\Theta(x)$ seien

$$\Theta(x), \Psi\Theta(x), \Psi^2\Theta(x), \ldots \Psi^{n-1}\Theta(x).$$

Ist dann $\psi(x)$ eine Transformation, die $\Theta(x)$ in $\Psi\Theta(x)$ überführt, also

$$\Theta\psi(x) = \Psi\Theta(x),$$

so folgt, dafs allgemein

$$\Theta\psi^i(x) = \Psi^i\Theta(x)$$

ist, und daher

$$x, \psi(x), \psi^2(x), \ldots \psi^{n-1}(x)$$

ein vollständiges System konjugierter Transformationen von G in Bezug auf H darstellen. Wir wollen nun annehmen, dafs n eine zusammengesetzte Zahl sei, p eine in ihr enthaltene Primzahl, so dafs

$$n = p\,n'$$

gesetzt werden kann. Setzen wir nun

$$\vartheta(x) = \psi^p(x),$$

so sind die Transformationen

$$x,\ \vartheta(x),\ \vartheta^2(x), \ldots \vartheta^{n'-1}(x)$$

konjugiert in Bezug auf H. Daher sind die Transformationen

$$\begin{array}{lll} x, & \varphi_1(x), & \ldots \varphi_{r-1}(x) \\ \vartheta(x), & \varphi_1\vartheta(x), & \ldots \varphi_{r-1}\vartheta(x) \\ \vdots & & \\ \vartheta^{n'-1}(x), & \varphi_1\vartheta^{n'-1}(x), & \ldots \varphi_{r-1}\vartheta^{n'-1}(x) \end{array}$$

von einander verschieden. Es läfst sich aber weiter zeigen, dafs sie eine Gruppe G' bilden. Setzt man nämlich zwei beliebige von ihnen $\varphi\vartheta^k$ und $\varphi'\vartheta^{k'}$ zusammen, so ist zunächst zu beachten, dafs $\vartheta^k\varphi'\bar{\vartheta}^k$ in H enthalten ist. Weil nämlich H eine ausgezeichnete Gruppe in G ist, so ist $\bar{\psi}\varphi_i\psi$ in H enthalten, folglich auch $\bar{\psi}\bar{\psi}\varphi_i\psi\psi = \bar{\psi}^2\varphi_i\psi^2$, allgemein $\bar{\psi}^h\varphi_i\psi^h$ und ebenso $\psi^h\varphi_i\bar{\psi}^h$. Wir können also, da $\vartheta = \psi^p$ war, $\vartheta^k\varphi'\bar{\vartheta}^k = \varphi''$ setzen, dann wird

$$\varphi\vartheta^k\varphi'\bar{\vartheta}^{k'} = \varphi\vartheta^k\varphi'\bar{\vartheta}^k\vartheta^{k+k'} = \varphi\varphi''\vartheta^{k+k'},$$

also eine in der Tabelle wieder enthaltene Transformation. Die Gruppe G' ist in G als ausgezeichnete enthalten, wie man leicht zeigt, und hat in Bezug auf sie den Index p, andrerseits enthält sie auch H als ausgezeichnete Untergruppe. Wir erhalten somit den Satz:

Zerfällt eine Galoissche Gleichung $P(x) = 0$ mit der Gruppe G durch Adjunktion der Wurzel einer cyklischen Gleichung, so kann man zu G eine in ihr ausgezeichnete Untergruppe G' angeben, deren Index in Bezug auf sie eine Primzahl p ist.

Man kann hier noch hinzufügen, dafs dann eine Zerfällung gewifs auch durch eine cyklische Gleichung vom Grade p erreicht werden kann Der Satz läfst sich aber sofort umkehren. Denn wenn G eine ausgezeichnete Untergruppe G' vom Primzahlindex hat, so läfst sie sich durch Adjunktion einer Galoisschen Gleichung vom Grade p reduzieren; eine Galoissche Funktion vom Primzahlgrade ist aber stets eine cyklische.

Fassen wir alles zusammen, was wir über Reduktion Galoisscher Gleichungen durch cyklische gefunden haben, so erhalten wir den Satz:

Die hinreichende und notwendige Bedingung dafür, dafs eine Galoissche Gleichung durch Wurzeln cyklischer Gleichungen reduziert werden kann, besteht darin, dafs ihre Gruppe G eine ausgezeichnete Untergruppe H vom Primzahlindex enthält.

Soll nun eine algebraische Gleichung durch Wurzeln cyklischer Gleichungen gelöst werden können, so mufs sich durch deren Adjunktion aus ihr ein Linearfaktor absondern lassen. Ein solcher ist aber dadurch charakterisiert, dafs er nur die identische Transformation zuläfst. Folglich mufs sich die ursprüngliche Gruppe auf die identische Transformation reduzieren lassen. Wir erhalten somit:

Die hinreichende und notwendige Bedingung dafür, dafs eine algebraische Gleichung durch Adjunktion cyklischer Körper lösbar ist, besteht darin, dafs die zugehörige Galoissche Funktion dadurch in lauter Linearfaktoren zerfällt, und dafs man daher zu ihrer Gruppe G eine Reihe von Gruppen

$$G, G_1, G_2, \ldots G_s$$

so bestimmen kann, dafs jede in der vorhergehenden als ausgezeichnete Untergruppe enthalten ist und in Bezug auf sie eine Primzahl als Index hat, und dafs diese Reihe bis zur identischen Transformation fortgeführt werden kann.

---

§ 107. Anwendung auf die Kreisteilungsgleichungen.

Schon in § 103 haben wir darauf hingewiesen, dafs die Primfunktion

$$P(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \ldots + x + 1$$

rationale Transformationen von der Gestalt

$$\varphi(x) = x^n$$

besitzt, wenn $(p, n) = 1$ ist, so dafs wir, indem wir n die Werte $1, 2, \ldots p - 1$ erteilen, im ganzen $(p - 1)$ inkongruente Transformationen

$$x, x^2, \ldots x^{p-2}, x^{p-1}$$

erhalten. Da deren Anzahl gleich dem Grade von $P(x)$ ist, so ist $P(x)$ eine Galoissche Funktion, und die Transformationen bilden daher eine Gruppe.

Um die Eigenschaften dieser Gruppe zu untersuchen, gehen wir von folgender Bemerkung aus. Sind

$$\varphi(x) = x^a, \quad \psi(x) = x^b$$

zwei Transformationen, so ist

$$\varphi\psi(x) = \psi\varphi(x) = x^{ab} \equiv x^c \mod P(x),$$

wenn $ab \equiv c \mod p$ ist. Zeigt sich hieran schon, dafs die Transformationen vertauschbar sind, also der Körper ein Abelscher ist, so ergiebt sich weiter, dafs die Exponenten a und b, die im verkürzten Restsystem nach dem Modul p enthalten sind, bei der Zusammensetzung zweier Transformationen multipliziert werden, wobei noch eine Reduktion auf $c \equiv ab \mod p$ erfolgen kann. Jeder Gruppe von Transformationen mufs also eine Gruppe des verkürzten Restsystem nach dem Modul p entsprechen und umgekehrt sich aus dieser Gruppe eine Gruppe von Transformationen ableiten lassen. Nun haben wir aber alle Gruppen des verkürzten Restsystems schon in § 71 untersucht und dabei gefunden, dafs die ganze Gruppe sich aus einem einzigen Grundelement g herleiten läfst, das als primitive Wurzel der Primzahl p definiert ist. Daher lassen sich auch die sämtlichen Transformationen der Gruppe aus der Transformation

$$\vartheta(x) = x^g$$

herleiten. Es ergiebt sich somit:

P(x) ist eine cyklische Funktion mit den Transformationen

$$x, \vartheta(x) = x^g, \vartheta^2(x) = x^{g^2}, \ldots \vartheta^{p-2}(x) = x^{g^{p-2}},$$

wenn g eine primitive Wurzel der Primzahl p bedeutet.

Die zu jedem Teiler von $(p-1)$ existierenden Untergruppen lassen sich dann auf folgende Weise darstellen: Es sei $p-1=ef$, so hat die Gruppe G_f von der Ordnung f die folgenden Transformationen:

$$x,\ x^{g^e},\ x^{g^{2e}},\ \ldots\ x^{g^{(f-1)e}}.$$

Das vollständige System der konjugierten Transformationen wird dargestellt durch

$$x,\ x^{g},\ x^{g^2},\ \ldots\ x^{g^{e-1}}.$$

Man kann die Gruppe G_f auch noch in anderer Weise zur Darstellung bringen, wie ebenfalls leicht bewiesen werden kann: Versteht man unter

$$a_0\equiv 1,\ a_1,\ a_2,\ \ldots\ a_{f-1}$$

die sämtlichen Lösungen der Kongruenz

$$x^f\equiv 1 \mod p,$$

so sind

$$x^{a_0}\equiv x,\ x^{a_1},\ x^{a_2},\ \ldots\ x^{a_{f-1}}$$

die Transformationen von G_f. Ist h eine primitive Wurzel derselben Kongruenz, so geht aus der Darstellung

$$x,\ x^{h},\ x^{h^2},\ \ldots\ x^{h^{f-1}}$$

die cyklische Natur von G_f mehr hervor.

Um eine primitive Gröfse des zu G_f gehörigen cyklischen Körpers zu bestimmen, gehen wir von der Bemerkung aus, dafs man statt des Fundamentalsystems

$$x,\ x^2,\ \ldots\ x^{p-1}$$

für den ganzen Körper auch das folgende zu Grunde legen kann

$$x,\ x^{g},\ x^{g^2},\ \ldots\ x^{g^{p-2}}.$$

Jede ganze Funktion $f(x)$ läfst sich also darstellen in der Form

$$f(x)=\sum_i u_i\, x^{g^i}. \qquad (i=0,1,\ldots p-2)$$

Soll nun $f(x)$ durch die Transformation x^{g^e} nicht geändert werden, so ergiebt sich, da

$$f(x^{g^e}) = \sum_i u_i\, x^{g^{e+i}} \qquad (i = 0, 1, \ldots p-2)$$

wird, daſs allgemein

$$u_i = u_{e+i}$$

sein muſs, so daſs $f(x)$ die Form annimmt

$$f(x) = \sum_i u_i \sum_k x^{g^{i+ek}}, \qquad \begin{pmatrix} i = 0, 1, \ldots e-1; \\ k = 0, 1, \ldots f-1 \end{pmatrix}$$

oder wenn wir

$$\Phi_f(x) = \sum_k x^{g^{ek}} \qquad (k = 0, 1, \ldots f-1)$$

setzen,

$$f(x) = \sum_i u_i\, \Phi_f(x^{g^i}). \qquad (i = 0, 1, \ldots e-1)$$

Es ergiebt sich also, daſs die e Funktionen

$$\Phi_f(x),\ \Phi_f(x^g), \ldots \Phi_f(x^{g^{e-1}})$$

ein Fundamentalsystem des zur Gruppe G_f gehörigen cyklischen Körpers sind, da zwischen ihnen keine lineare homogene Relation mit rationalen Koeffizienten bestehen kann. Die Funktionen sind konjugiert und alle von einander verschieden. Wir können daher eine von ihnen als primitives Element des Körpers zu Grunde legen. Um die Beziehungen zu ermitteln, die zwischen ihnen bestehen, also die cyklische Gleichung aufzustellen, der sie genügen, und alle rational durch eine einzige auszudrücken, ist es nur nötig, das Produkt je zweier linear und homogen durch alle darstellen zu können.

Ist nun

$$\Phi_f(x) = x^{a_0} + x^{a_1} + \ldots + x^{a_{f-1}},$$

so ergiebt sich

$$\Phi_f(x^i)\,\Phi_f(x^k) = \sum_s x^{a_s i}\,\Phi_f(x^k) = \sum_s x^{a_s i}\,\Phi_f(x^{a_s k})$$

$$= \sum_{s,t} x^{a_s i}\, x^{a_s a_t k} = \sum_{s,t} x^{a_s(i + a_t k)} = \sum_t \Phi_f(x^{i + a_t k}),$$

$$(s, t = 0, 1, \ldots f-1)$$

wobei berücksichtigt wurde, daſs $\Phi_f(x^k) = \Phi_f(x^{a_s k})$ ist. Ist also

$$\Phi_f(x^i) = x^i + x^{i_1} + \ldots + x^{i_{f-1}}$$

$$\Phi_f(x^k) = x^k + x^{k_1} + \ldots + x^{k_{f-1}},$$

so wird die gesuchte Darstellung durch die Formel

$$\Phi_f(x^i)\,\Phi_f(x^k) = \sum_t \Phi_f(x^{i+k_t}) \qquad (t = 0, 1, \ldots f-1)$$

gegeben.

Wir wollen die Reduktion auf cyklische Gleichungen nur an einigen Beispielen erläutern:

1) **$p = 13$.** Da $p - 1 = 12 = 2^2 . 3$ ist, so läfst sich die Auflösung der Gleichung auf eine cyklische Gleichung dritten Grades und zwei quadratische zurückführen. Wir beginnen mit der Aufstellung der Gleichung dritten Grades und bemerken, dafs die Gruppe vierter Ordnung im verkürzten Restsystem aus den Resten 1, 5, 8, 12 besteht, und dafs, da 2 eine primitive Wurzel von 13 ist, die konjugierten Reste in Bezug hierauf durch 1, 2, 4 dargestellt werden können. Wir haben so

$$\begin{aligned}
\Phi_4(x) &= x + x^5 + x^8 + x^{12} \\
\Phi_4(x^2) &= x^2 + x^3 + x^{10} + x^{11} \\
\Phi_4(x^4) &= x^4 + x^6 + x^7 + x^9.
\end{aligned}$$

Bilden wir nun die Produkte

$$\begin{aligned}
\Phi_4(x)^2 &= -4\,\Phi_4(x) - 3\,\Phi_4(x^2) - 2\,\Phi_4(x^4) \\
\Phi_4(x)\,\Phi_4(x^2) &= \Phi_4(x) + 2\,\Phi_4(x^2) + \Phi_4(x^4) \\
\Phi_4(x)\,\Phi_4(x^4) &= 2\,\Phi_4(x) + \Phi_4(x^2) + \Phi_4(x^4),
\end{aligned}$$

so erhalten wir durch Elimination von $\Phi_4(x^2)$, $\Phi_4(x^4)$ die folgende Gleichung für die konjugierten Funktionen

$$Q(y) = \begin{vmatrix} 4+y & 3 & 2 \\ 1 & 2-y & 1 \\ 2 & 1 & 1-y \end{vmatrix} = y^3 + y^2 - 4y + 1 = 0,$$

und wenn wir $\Phi_4(x^2)$ und $\Phi_4(x^4)$ durch $\Phi_4(x)$ ausdrücken, so finden wir

$$\begin{aligned}
\Phi_4(x^2) &= -\Phi_4(x)^2 - 2\,\Phi_4(x) + 2 \\
\Phi_4(x^4) &= \Phi_4(x)^2 + \Phi_4(x) - 3.
\end{aligned}$$

Man bestätigt leicht, dafs die Funktionen

$$\psi(y) = -y^2 - 2y + 2, \quad \psi^2(y) = y^2 + y - 3$$

rationale Transformationen von $Q(y)$ sind.

Um nun die erste quadratische Gleichung zu finden, bemerken wir, dafs

$$\Phi_2(x) = x + x^{12}, \quad \Phi_2(x^5) = x^5 + x^8$$

also

$$\Phi_2(x) + \Phi_2(x^5) = \Phi_4(x)$$
$$\Phi_2(x)\,\Phi_2(x^5) = \Phi_4(x^4) = \Phi_4(x)^2 + \Phi_4(x) - 3$$

ist. Daher genügen $\Phi_2(x)$ und $\Phi_2(x^5)$ der Gleichung

$$z^2 - \Phi_4(x)\,z + \Phi_4(x)^2 + \Phi_4(x) - 3 = 0.$$

Endlich genügen x und x^{12}, da

$$x + x^{12} = \Phi_2(x), \quad x\,x^{12} = 1$$

ist, der quadratischen Gleichung

$$x^2 - \Phi_2(x)\,x + 1 = 0.$$

2) $p = 17$. Da $p - 1 = 16 = 2^4$ ist, so läfst sich die Gleichung auf vier quadratische zurückführen.

Zunächst wird (§ 72)

$$\Phi_8(x) \;= x + x^2 + x^4 + x^8 + x^9 + x^{13} + x^{15} + x^{16}$$
$$\Phi_8(x^3) = -1 - \Phi_8(x)$$
$$\Phi_8(x)^2 = 3\,\Phi_8(x) + 4\,\Phi_8(x^3) + 8 = -\Phi_8(x) + 4,$$

so dafs $\Phi_8(x)$ wie $\Phi_8(x^3)$ der Gleichung

$$y^2 + y - 4 = 0$$

genügen. Darauf betrachten wir

$$\Phi_4(x) = x + x^4 + x^{13} + x^{16}$$

sowie die konjugierten Funktionen

$$\Phi_4(x^2) = x^2 + x^8 + x^9 + x^{15}$$
$$\Phi_4(x^3) = x^3 + x^5 + x^{12} + x^{14}$$
$$\Phi_4(x^6) = x^6 + x^7 + x^{10} + x^{11}.$$

Da $\Phi_4(x)$ eine primitive Gröfse des Körpers ist, der durch Adjunktion von $\Phi_8(x)$ erhalten wurde, so müssen die Gröfsen

$$1, \ \Phi_8(x), \ \Phi_4(x), \ \Phi_8(x)\,\Phi_4(x)$$

als Fundamentalsystem dienen können. Nun ist

$$
\begin{aligned}
-1 &= \Phi_4(x) + \Phi_4(x^2) + \Phi_4(x^3) + \Phi_4(x^6) \\
\Phi_8(x) &= \Phi_4(x) + \Phi_4(x^2) \\
\Phi_4(x)\,\Phi_8(x) &= 3 + \Phi_4(x^2) + 2\,\Phi_4(x^3),
\end{aligned}
$$

und daraus folgt

$$
\begin{aligned}
\Phi_4(x^2) &= \Phi_8(x) - \Phi_4(x) \\
\Phi_4(x^3) &= \frac{1}{2}[-3 + \Phi_4(x) - \Phi_8(x) + \Phi_4(x)\,\Phi_8(x)] \\
\Phi_4(x^6) &= \frac{1}{2}[1 - \Phi_4(x) - \Phi_8(x) - \Phi_4(x)\,\Phi_8(x)].
\end{aligned}
$$

Da nun

$$\Phi_4(x)^2 = 4 + \Phi_4(x^2) + 2\,\Phi_4(x^3) = 1 + \Phi_4(x)\,\Phi_8(x)$$

ist, so genügt $\Phi_4(x)$ der Gleichung

$$z^2 - \Phi_8(x)\,z - 1 = 0,$$

deren andere Wurzel, weil $\Phi_8(x)$ durch x^2 ungeändert bleibt, $\Phi_4(x^2)$ sein mufs.

Weiter betrachten wir

$$\Phi_2(x) = x + x^{16},$$

bilden

$$
\begin{aligned}
\Phi_2(x) + \Phi_2(x^4) &= \Phi_4(x) \\
\Phi_2(x)\,\Phi_2(x^4) &= \Phi_2(x^5) + \Phi_2(x^3) = \Phi_4(x^3)
\end{aligned}
$$

und erhalten somit für $\Phi_2(x)$ und $\Phi_2(x^4)$ die quadratische Gleichung

$$u^2 - \Phi_4(x)\,u + \frac{1}{2}[-3 + \Phi_4(x) - \Phi_8(x) + \Phi_4(x)\,\Phi_8(x)] = 0.$$

Endlich gilt für x und x^{16} die Gleichung

$$x^2 - \Phi_2(x)\,x + 1 = 0.$$

Die ursprüngliche Gleichung ist somit auf die folgende Kette von quadratischen Gleichungen zurückgeführt:

$$
\begin{gathered}
x^2 - u\,x + 1 = 0 \\
u^2 - z\,u + \frac{1}{2}[-3 + z - y + z\,y] = 0 \\
z^2 - y\,z - 1 = 0,\quad y^2 + y - 4 = 0.
\end{gathered}
$$

Es ist leicht, diese aufzulösen und so x durch lauter Quadratwurzeln darzustellen.

---

§ 108. Eigenschaften der binomischen Gleichungen.

Wie wir bereits in § 98 angedeutet haben, ist es eine der Hauptaufgaben dieses Abschnittes, die hinreichenden und notwendigen Bedingungen dafür festzustellen, dafs sich die Auflösung einer algebraischen Gleichung auf die einer Reihe von reinen oder binomischen Gleichungen zurückführen läfst. Wir müssen uns daher zunächst etwas näher mit diesen beschäftigen. Eine solche Gleichung hat die Form

$$R(x) = x^p - a = 0,$$

wo p als Primzahl vorausgesetzt wird, und a keine p te Potenz einer rationalen Gröfse sein soll, in welchem Falle die Gleichung ja eine rationale Wurzel besitzen würde. Wir wollen uns nun bei unserer Untersuchung auf den Standpunkt stellen, dafs wir die primitiven Einheitswurzeln, d. h. die Gröfse α, die der Gleichung

$$\frac{y^p - 1}{y - 1} = y^{p-1} + y^{p-2} + \dots + y + 1 = 0$$

genügt, dem Rationalitätsbereich adjungieren. Wie leicht zu erkennen ist, sind dann

$$x, \alpha x, \alpha^2 x, \dots \alpha^{p-1} x$$

Transformationen von R(x). Man zeigt auf einfache Weise, dafs sie eine cyklische Gruppe bilden von der Ordnung p, die durch Iteration der Transformation

$$\varphi(x) = \alpha x$$

hervorgeht, denn es wird

$$\varphi^2(x) = \alpha^2 x, \ \varphi^3(x) = \alpha^3 x, \dots \varphi^{p-1}(x) = \alpha^{p-1} x$$

und $\varphi^p(x) = \alpha^p x = x$. Wenn wir nun noch zeigen können, dafs R(x) in dem vorausgesetzten Rationalitätsbereich irreduktibel ist, so folgt, dafs R(x) eine Galoissche Funktion und zwar eine cyklische ist. Dies läfst sich aber sofort daraus erschliefsen, dafs ein Primteiler P(x) von R(x), wenn er nicht linear ist, eine nicht identische Transformation der Gruppe besitzt; denn dann wird die Gleichung P(x) = 0 durch mindestens zwei verschiedene p te Wurzeln aus a befriedigt, die sich um eine Einheitswurzel als Faktor unterscheiden. Aus einer nicht identischen Transformation lassen

sich aber alle $p - 1$ übrigen durch Zusammensetzung bilden, und diese müssen also auch Transformationen von $P(x)$ sein. $P(x)$ hat also mindestens den Grad p, und es folgt somit:

Wenn $R(x)$ im Rationalitätsbereich der $p$ten Einheitswurzeln nicht in lauter Linearfaktoren zerlegbar ist, von denen dann einer rational sein mufs, so ist $R(x)$ in ihm eine cyklische Primfunktion.

Dieser Satz läfst sich in folgender Weise umkehren: Die Auflösung jeder cyklischen Gleichung $P(x) = 0$ vom Primzahlgrade p läfst sich unter Adjunktion $p$ter Einheitswurzeln auf eine reine Gleichung zurückführen.

Sind nämlich die rationalen Transformationen von $P(x)$ dargestellt in der Form

$$x, \theta(x), \theta^2(x), \ldots \theta^{p-1}(x),$$

so dafs $\theta^p(x) \equiv x \mod P(x)$ ist, und bildet man dann die sogenannte Lagrangesche Resolvente

$$\Phi(x, \alpha) = x + \alpha\,\theta(x) + \alpha^2\,\theta^2(x) + \ldots + \alpha^{p-1}\,\theta^{p-1}(x),$$

so wird

$$\Phi(\theta(x), \alpha) = \theta(x) + \alpha\theta^2(x) + \alpha^2\,\theta^3(x) + \ldots + \alpha^{p-1}x$$
$$= \alpha^{-1}\,\Phi(x, \alpha)$$

und allgemein

$$\Phi(\theta^i(x), \alpha) \equiv \alpha^{-i}\,\Phi(x, \alpha) \mod P(x).$$

Hieraus ergiebt sich sofort durch Multiplikation

$$[\Phi(x, \alpha)]^p = \Phi(x, \alpha)\,\Phi(\theta(x), \alpha) \ldots \Phi(\theta^{p-1}(x), \alpha).$$

Auf der rechten Seite steht aber eine Funktion von x, die durch keine Transformation ihren Wert ändert und somit rational bekannt ist, wenn man von den Einheitswurzeln α absieht. Da $\Phi(x, \alpha)$ durch jede Transformation von $P(x)$ seinen Wert ändert, so ist es eine primitive Gröfse im cyklischen Körper $P(x) = 0$, und daher läfst sich x rational durch $\Phi(x, \alpha)$ ausdrücken, das als Wurzel einer binomischen Gleichung bestimmt ist.

§ 109. Bedingung für die Auflösbarkeit der algebraischen Gleichungen durch Wurzelgröfsen.

Nachdem wir erkannt haben, dafs die reinen Gleichungen vom Grade p im Rationalitätsbereich der pten Einheitswurzeln cyklische Gleichungen sind, und umgekehrt die Wurzeln dieser rational durch die Wurzeln der reinen Gleichung dargestellt werden können, können wir sofort den Schlufssatz des § 106 zur Anwendung bringen, der uns auch daher die hinreichenden und notwendigen Bedingungen dafür angiebt, wann eine Gleichung durch Wurzelgröfsen lösbar ist. Hierbei wird aber immer vorausgesetzt, dafs man die nötigen Einheitswurzeln adjungieren darf. Wird dadurch keine der Gruppen geändert, so ist diese Adjunktion ohne Belang. Es kann aber sehr wohl vorkommen, dafs hierdurch eine Erniedrigung des Galoisschen Körpers eintritt. Da aber alle im Körper der Einheitswurzeln enthaltenen Körper cyklisch sind, so tritt eine Erniedrigung nur ein, wenn eine der Gruppen eine ausgezeichnete Untergruppe vom Primzahlindex besitzt, der im Grade des Körpers enthalten ist. Man erkennt daraus, dafs die formulierte Bedingung bei Bestand bleibt. Wir können daher allgemein behaupten:

Die hinreichende und notwendige Bedingung dafür, dafs sich eine algebraische Gleichung durch Wurzelgröfsen lösen läfst, besteht darin, dafs man zur Gruppe G der rationalen Transformationen ihrer Galoisschen Resolvente eine Reihe von Gruppen

$$G, G_1, G_2, \dots G_s$$

bestimmen kann, von denen jede folgende in der vorhergehenden als ausgezeichnete Untergruppe enthalten ist und in Bezug auf sie eine Primzahl als Index besitzt, und von denen die letzte Gruppe eine Primzahl als Ordnung hat.

Ist der Galoissche Körper ein cyklischer, so ist diese Bedingung stets erfüllt; denn wir haben in § 106 bewiesen, dafs es zu jedem Teiler der Ordnung der Gruppe eine ausgezeichnete cyklische Untergruppe giebt, und daraus folgt dann, dafs man auch eine solche Untergruppe bestimmen kann, deren Index eine Primzahl ist. Allgemein

läfst sich dies auf alle Abelschen Körper übertragen, doch würde es uns zu weit führen, den Beweis hierfür darzulegen. Wir wollen nur noch darauf hinweisen, dafs die in § 107 betrachteten Kreisteilungsgleichungen als cyklische Gleichungen zu den durch Wurzelgröfsen auflösbaren gehören. Wir haben ihre Reduktion auf einfachere cyklische Gleichungen schon dort an zwei Beispielen durchgeführt, und es würde ein Leichtes sein, durch Einführung der Lagrangeschen Resolventen den cyklischen Gleichungen die Form reiner Gleichungen zu geben.

---

§ 110. Permutationsgruppe einer Gleichung.

Es sei $F(x) = 0$ eine irreduktible algebraische Gleichung $n$ten Grades mit den Wurzeln $x_0, x_1, x_2, \ldots x_{n-1}$, t die Galoissche Resolvente und $P(t) = 0$ die zugehörige Gleichung vom Grade m mit der Gruppe G von rationalen Transformationen. Nehmen wir an, dafs eine Wurzel von $F(x) = 0$ durch

$$x_0 = \chi(t)$$

dargestellt wird, so mufs $\chi(t)$ durch die Transformationen

$$\varphi_0(t) = t, \varphi_1(t), \ldots \varphi_{r-1}(t)$$

einer Gruppe H von der Ordnung $r = \frac{m}{n}$ nicht geändert werden, durch jede andere Transformation aus G dagegen in konjugierte Werte übergehen. Sind dann

$$t, \psi_1(t), \psi_2(t), \ldots \psi_{n-1}(t)$$

ein vollständiges System konjugierter Transformationen von G in Bezug auf H, so können die sämtlichen Wurzeln der Gleichung $F(x) = 0$ dargestellt werden durch

$$x_0 = \chi(t), x_1 = \chi\psi_1(t), \ldots x_{n-1} = \chi\psi_{n-1}(t).$$

Bei Anwendung irgend einer Transformation $\varrho(t)$ aus der Gruppe G entsteht wieder ein System konjugierter Werte, da, wie leicht einzusehen ist, stets zwei verschiedene konjugierte Funktionen wieder in zwei verschiedene übergehen. Daher können wir behaupten:

Jeder Transformation von G entspricht eine Permutation der Wurzeln $x_0, x_1, \ldots x_{n-1}$.

Dieses gilt überhaupt, wenn wir unter $x_0, x_1, \ldots x_{n-1}$ irgend ein System konjugierter Werte einer Funktion von t verstehen. Nehmen wir nun aber hinzu, daſs t die Galoissche Resolvente ist, also rational durch $x_0, x_1, \ldots x_{n-1}$ dargestellt werden kann in der Form

$$t \equiv \theta(\chi(t), \chi\psi_1(t), \ldots \chi\psi_{n-1}(t)) \bmod P(t),$$

so folgt durch Anwendung zweier verschiedener Transformationen $\varrho(t)$ und $\varrho'(t)$ aus G, daſs

$$\varrho(t) \equiv \theta(\chi\varrho(t), \chi\psi_1\varrho(t), \ldots \chi\psi_{n-1}\varrho(t)) \bmod P(t)$$
$$\varrho'(t) \equiv \theta(\chi\varrho'(t), \chi\psi_1\varrho'(t), \ldots \chi\psi_{n-1}\varrho'(t)) \bmod P(t)$$

ist und daraus der Satz:

Verschiedenen Transformationen $\varrho(t)$ und $\varrho'(t)$ aus G entsprechen auch verschiedene Permutationen P und P'.

Wäre dies nämlich nicht der Fall, so wären die rechten Seiten der Kongruenzen nicht verschieden (nach P(t) als Modul) und daher auch die linken Seiten nicht, die der Annahme nach verschiedene Transformationen bedeuten sollen.

Die Permutation P können wir in folgender Weise darstellen

$$P = \begin{pmatrix} \chi(t), & \chi\psi_1(t), & \ldots \chi\psi_{n-1}(t) \\ \chi\varrho(t), & \chi\psi_1\varrho(t), & \ldots \chi\psi_{n-1}\varrho(t) \end{pmatrix}$$

ähnlich so P'. Wir können aber P' zum Zwecke der Zusammensetzung mit P auch in folgender Form schreiben

$$P' = \begin{pmatrix} \chi\varrho(t), & \chi\psi_1\varrho(t), & \ldots \chi\psi_{n-1}\varrho(t) \\ \chi\varrho\varrho'(t), & \chi\psi_1\varrho\varrho'(t), & \ldots \chi\psi_{n-1}\varrho\varrho'(t) \end{pmatrix}.$$

Dann folgt sofort

$$P\,P' = \begin{pmatrix} \chi(t), & \chi\psi_1(t), & \ldots \chi\psi_{n-1}(t) \\ \chi\varrho\varrho'(t), & \chi\psi_1\varrho\varrho'(t), & \ldots \chi\psi_{n-1}\varrho\varrho'(t) \end{pmatrix}.$$

Entsprechen also den Transformationen $\varrho(t)$ und $\varrho'(t)$ der Gruppe G die Permutationen

P und P' der Wurzeln, so entspricht der Transformation $\varrho\varrho'(t)$ die Permutation PP'.

Hieraus folgt nun sofort, daſs die den Transformationen entsprechenden Permutationen ebenso wie diese eine Gruppe bilden. Zwei Gruppen, die aus gleichvielen Elementen bestehen, die einander umkehrbar eindeutig so zugeordnet werden können, daſs diese Zuordnung auch bei der Zusammensetzung bestehen bleibt, nennt man isomorphe Gruppen. Wir können daher den Satz aussprechen:

Für jede algebraische Gleichung giebt es eine Permutationsgruppe der Wurzeln, die der Gruppe der rationalen Transformationen der Galoisschen Resolvente isomorph ist.

Diese Gruppe heiſst die Galoissche Permutationsgruppe der Gleichung.

Wir haben früher in § 99 gefunden, daſs alle symmetrischen Funktionen der Wurzeln rational durch die Koeffizienten dargestellt werden können. Wir sind jetzt imstande, diesen Satz zu verallgemeinern, auch für den Fall, daſs man beliebige Irrationalitäten adjungiert hat, und daſs zwischen den Wurzeln beliebige Beziehungen bestehen. Es besteht nämlich der Satz:

Die hinreichende und notwendige Bedingung dafür, daſs eine rationale Funktion der Wurzeln rational bekannt ist, besteht darin, daſs sie ihren Wert nicht ändert, wenn sie den Permutationen der Gruppe in Bezug auf den vorliegenden Rationalitätsbereich unterworfen wird.

Jede rationale Funktion der Wurzel läſst sich nämlich als rationale Funktion f(t) der Galoisschen Resolvente t darstellen. Bleibt sie durch alle Permutationen der Gruppe ungeändert, so läſst sie als Funktion von t betrachtet alle rationalen Transformationen von t zu, bildet also einen Körper ersten Grades, d. h. ist eine Gröſse a des Rationalitätsbereiches. Und umgekehrt hat sie den rationalen Wert a, so kann man auf $f(t) = a$ die Transformationen von P(t) anwenden, von denen eine jede einer Permutation der Gruppe entspricht.

Da die Permutationsgruppe der Transformationsgruppe

isomorph ist, so entspricht jeder Untergruppe der einen auch eine solche der andern und umgekehrt. Einer Untergruppe, die in der einen ausgezeichnet enthalten ist, entspricht eine solche bei der andern. Man kann also die eine Gruppe genau so zerlegen wie die andere, und da diese Art der Zerlegung, nicht die spezielle Form der Transformationen oder Permutationen, bei der algebraischen Auflösung der Gleichungen allein eine Rolle spielt, so kann die eine Gruppe völlig die andere vertreten. Bei den Kreisteilungsgleichungen ist es vorteilhaft, die Transformationsgruppen in den Vordergrund zu stellen. Sobald es sich jedoch um allgemeine Gleichungen handelt, zwischen deren Wurzeln keine Beziehungen obwalten, geht man besser von der Permutationsgruppe aus, die in diesem Falle aus allen möglichen Permutationen der Wurzeln besteht. Wir haben nun bereits in § 42 gesehen, dafs die vollständige Gruppe die Alterngruppe ausgezeichnet enthält, dafs diese aber für mehr als vier Elemente eine einfache Gruppe darstellt, deren Grad eine zusammengesetzte Zahl ist. Hieraus ergiebt sich dann sofort, dafs es unmöglich ist, allgemeine Gleichungen vom höheren als vierten Grade durch Wurzelgröfsen zu lösen.

Dagegen sind die allgemeinen Gleichungen niederen Grades lösbar. Die Gleichungen zweiten Grades besitzen als Permutationsgruppe eine einfache Transposition aufser der identischen Permutation, sind also durch Quadratwurzeln lösbar. Da bei ihnen wie bei den kubischen Gleichungen und den biquadratischen in der vollständigen Gruppe die Alterngruppe mit dem Index 2 ausgezeichnete Untergruppe ist, so wird diese die Gruppe der Gleichung, sobald man eine geeignete zu ihnen gehörige Quadratwurzel adjungiert. Bei den kubischen Gleichungen ist die Alterngruppe eine einfache Gruppe von der Ordnung 3, daher mufs die Adjunktion einer zu ihr gehörigen Funktion der Wurzeln sodann die Gleichung zu einer cyklischen machen. Bei den biquadratischen Gleichungen enthält die Alterngruppe zunächst eine ausgezeichnete Untergruppe mit dem Index 3, und aus dieser läfst sich (auf drei Weisen) eine Gruppe vom Index 2 absondern, wobei eine einfache Gruppe von der Ordnung 2 übrig bleibt. Demnach hat man bei der Auflösung der biquadratischen Gleichungen der Reihe nach erst

eine Quadratwurzel, darauf eine Kubikwurzel und schliefslich noch zwei Quadratwurzeln auszuziehen. Die genaue Durchführung dieser Auflösungsmethoden wollen wir nun in den beiden folgenden Paragraphen betrachten.

---

§. 111. Auflösung der kubischen Gleichungen.

Um nun darzulegen, wie sich aus der in den vorhergehenden Paragraphen entwickelten Theorie eine naturgemäfse Auflösung der kubischen Gleichung

$$F(x) = x^3 - c_1 x^2 + c_2 x - c_3 = 0$$

ergiebt, haben wir zunächst bezüglich der Permutationsgruppe auf die Darstellung in § 39 hinzuweisen. Nennen wir die Wurzeln der kubischen Gleichung x_0, x_1, x_2, so ergiebt sich, dafs diese Gruppe aus den beiden Permutationen

$$S = (x_0, x_1, x_2), \quad T = (x_1, x_2)$$

durch wiederholte Zusammensetzung gebildet werden kann. Die einzige vorhandene ausgezeichnete Untergruppe ist die Alterngruppe, deren Permutationen durch Iteration von S erhalten werden, die also die Permutationen

$$1, S = (x_0, x_1, x_2), \quad S^2 = (x_0, x_2, x_1)$$

enthält. Da ihr Index gleich 2 ist, so wird sie die Gruppe der Gleichung, wenn man eine zu ihr gehörige Funktion adjungiert, die dann einer quadratischen Gleichung genügt. Soll diese quadratische Gleichung eine reine sein, so mufs die Funktion die Eigenschaft haben, das entgegengesetzte Vorzeichen anzunehmen, wenn man die Permutation T zur Anwendung bringt. Wir haben in § 101 die allgemeine Form dieser Funktion bestimmt und setzen nun

$$\varphi(x_0, x_1, x_2) = (x_0 - x_1)(x_0 - x_2)(x_1 - x_2).$$

Die quadratische Gleichung, der sie genügt, läfst sich leicht bilden, denn $\varphi(x_0, x_1, x_2)^2$ ist eine symmetrische Funktion von x_0, x_1, x_2, da sie weder durch S noch durch T eine Wertänderung erleidet, und als solche durch die symmetrischen Grundformen f_1, f_2, f_3 in der Gestalt

$$18 f_1 f_2 f_3 + f_1^2 f_2 - 4 f_1^3 f_3 - 4 f_2^3 - 27 f_3^2$$

darstellbar. Daher genügen $\varphi(x_0, x_1, x_2)$ und $\varphi(x_0, x_2, x_1) = -\varphi(x_0, x_1, x_2)$ der quadratischen Gleichung

$$y^2 - D = 0,$$

wo D die Discriminante der kubischen Gleichung darstellt.

Jetzt müssen wir eine Funktion suchen, deren Wert sich durch die Permutation S ändert. Eine solche ist x_0, was durch S, S^2 bezw. in x_1, x_2 übergeht. Adjungiert man also $\varphi(x_0, x_1, x_2)$, so wird die kubische Gleichung eine cyklische, und x_1, x_2 lassen sich rational durch x_0 ausdrücken. Man kann dies sofort auch daraus erkennen, dafs $\varphi(x_0, x_1, x_2)$ in der Form darstellbar ist (§ 101)

$$\varphi(x_0, x_1, x_2) = 6x_0^2 x_1 - 2c_1 x_0^2 - 4c_1 x_0 x_1 + 2(c_1^2 - c_2)x_0 + 2c_2 x_1 - c_1 c_2 + 3c_3,$$

denn diese Gleichung gestattet es sofort, x_1 durch $\varphi(x_0, x_1, x_2)$ und x_0 darzustellen. Nimmt man dann noch hinzu, dafs

$$x_2 = c_1 - x_1 - x_0$$

ist, so folgt auch eine Darstellung von x_2. Diese Ausdrücke sind allerdings gebrochene Funktionen. Wollte man eine Darstellung in Gestalt ganzer Funktionen erzielen, so brauchte man nur $\varphi(x_0, x_1, x_2)x_0$, $\varphi(x_0, x_1, x_2)x_0^2$ in reduzierter Form zu berechnen und würde dann sofort x_1 durch Auflösung erhalten. Dem Leser möge die Ausführung dieser Rechnung überlassen bleiben und nur noch bemerkt werden, dafs die Darstellung von x_1 als lineare gebrochene Funktion von x_0 eine besonders einfache Gestalt annimmt.

Für die algebraische Auflösung ist aber die Gröfse x_0 nicht geeignet, weil sie keiner reinen kubischen Gleichung genügt. Eine solche Funktion erhalten wir nach § 108, wenn wir eine primitive dritte Einheitswurzel α adjungieren d. h. eine Wurzel der Gleichung

$$\frac{x^3 - 1}{x - 1} = x^2 + x + 1 = 0$$

und die Lagrangesche Resolvente

$$\psi(x_0, x_1, x_2) = x_0 + \alpha x_1 + \alpha^2 x_2$$

bilden, die den Bedingungen

$$\psi_s = \alpha^2 \psi, \quad \psi_{s^2} = \alpha \psi$$

genügt und Wurzel einer reinen kubischen Gleichung ist. Man kann diese, sowie die rationalen Ausdrücke von x_0, x_1, x_2 durch φ und ψ auf rein methodischem Wege finden, wenn man die Ausdrücke

$$\varphi, \psi, \psi^2, \varphi \cdot \psi, \varphi \cdot \psi^2$$

reduziert, die zusammen mit der Einheit ein Fundamentalsystem für den kubischen Körper bilden. Es ist dem Leser sehr zu empfehlen, dies wirklich auszuführen, um dadurch eine völlig klare Einsicht in den algebraischen Auflösungsprozefs zu erhalten. Wir ziehen es jedoch im Interesse der Kürze vor, einen etwas andern Weg einzuschlagen. Man erhält zunächst

$$\begin{aligned}\psi(x_0, x_1, x_2)^3 &= x_0^3 + x_1^3 + x_2^3 \\ &+ 6 x_0 x_1 x_2 - 3(x_0 x_1^2 + x_1 x_2^2 + x_2 x_0^2) \\ &+ 3\alpha(x_0 - x_1)(x_0 - x_2)(x_1 - x_2).\end{aligned}$$

Die auftretende unsymmetrische Funktion $x_0 x_1^2 + x_1 x_2^2 + x_2 x_0^2$ läfst sich durch $\varphi(x_0, x_1, x_2)$ ausdrücken, wenn man beachtet, dafs

$$\begin{aligned}\varphi(x_0, x_1, x_2) &= (x_0^2 x_1 + x_1^2 x_2 + x_2^2 x_0) - (x_0 x_1^2 + x_1 x_2^2 + x_2 x_0^2) \\ f_1 f_2 - 3 f_3 &= (x_0^2 x_1 + x_1^2 x_2 + x_2^2 x_0) + (x_0 x_1^2 + x_1 x_2^2 + x_2 x_0^2)\end{aligned}$$

ist. Nimmt man nun noch hinzu, dafs

$$x_0^3 + x_1^3 + x_2^3 = f_1^3 - 3 f_1 f_2 + 3 f_3, \quad x_0 x_1 x_2 = f_3$$

ist, so findet man, dafs $\psi(x_0, x_1, x_2)$ der reinen kubischen Gleichung

$$z^3 = c_1^3 - \frac{9}{2} c_1 c_2 + \frac{27}{2} c_3 + \frac{3}{2}(1 + 2\alpha)\varphi(x_0, x_1, x_2)$$

genügt. Die andern Wurzeln dieser Gleichung sind dann $\psi_S = \alpha^2 \psi$ und $\psi_{S^2} = \alpha \psi$. Wendet man die Permutation T an, so findet man die kubische Gleichung, der die Gröfsen $\psi_T, \psi_{ST}, \psi_{S^2T}$ genügen; sie unterscheidet sich von der hingeschriebenen nur dadurch, dafs φ durch $\varphi_T = -\varphi$ ersetzt werden mufs.

Um nun zu den Wurzeln x_0, x_1, x_2 zu gelangen, beachten wir, dafs

$$\begin{aligned} x_0 + x_1 + x_2 &= c_1 \\ x_0 + \alpha\, x_1 + \alpha^2 x_2 &= \psi \\ x_0 + \alpha^2 x_1 + \alpha\, x_2 &= \psi_T \end{aligned}$$

ist, dann finden wir sofort

$$x_0 = \frac{1}{3}(c_1 + \psi + \psi_T)$$

$$x_1 = \frac{1}{3}(c_1 + \alpha^2 \psi + \alpha\, \psi_T)$$

$$x_2 = \frac{1}{3}(c_1 + \alpha\, \psi + \alpha^2 \psi_T).$$

Dies ist aber noch nicht die gewünschte Darstellung, da ψ_T noch nicht rational durch ψ ausgedrückt worden ist. Um diese zu erhalten, bemerke man, dafs das Produkt $\psi \cdot \psi_T$ weder durch S noch durch T seinem Werte nach geändert wird und daher eine symmetrische Funktion von x_0, x_1, x_2 ist. Man findet in der That

$$\psi \cdot \psi_T = x_0^2 + x_1^2 + x_2^2 - x_0 x_1 - x_0 x_2 - x_1 x_2 = f_1^2 - 3 f_2,$$

also

$$\psi_T = \frac{c_1^2 - 3 c_2}{\psi}.$$

§ 112. Auflösung der biquadratischen Gleichungen.

Um die Auflösung der allgemeinen biquadratischen Gleichung

$$F(x) = x^4 - c_1 x^3 + c_2 x^2 - c_3 x + c_4 = 0$$

zu entwickeln, haben wir zunächst an die Ergebnisse des § 41 zu erinnern, wo die allgemeine Permutationsgruppe von vier Elementen in Rücksicht auf die in ihr enthaltenen ausgezeichneten Untergruppen untersucht wurde. Nennen wir die Wurzeln der biquadratischen Gleichung x_0, x_1, x_2, x_3, so soll jetzt

$$S_1 = (x_0, x_1)(x_2, x_3), \quad S_2 = (x_0, x_2)(x_1, x_3)$$
$$T = (x_1, x_2, x_3), \quad U = (x_1, x_2)$$

angenommen werden. In der vollständigen Gruppe sind zwei ausgezeichnete Untergruppen, nämlich die Alterngruppe G und die Gruppe

$$H = \{1, S_1, S_2, S_1 S_2\}$$

enthalten, von denen die letztere sich wieder in zwei Gruppen

$$K_1 = (1, S_1), \; K_2 = (1, S_2)$$

zerlegen läfst.

Um die Auflösung durchzuführen, hat man zunächst eine zur Alterngruppe gehörende Funktion zu bilden. Eine solche ist

$$\varphi(x_0, x_1, x_2, x_3)$$
$$= (x_0 - x_1)(x_0 - x_2)(x_0 - x_3)(x_1 - x_2)(x_1 - x_3)(x_2 - x_3),$$

denn es wird

$$\varphi_{S_1} = \varphi_{S_2} = \varphi_T = \varphi.$$

Da

$$\varphi_U = -\varphi$$

wird, so genügt φ sowie $\varphi_U = -\varphi$ einer rein quadratischen Gleichung

$$y^2 - D = 0,$$

wo D, wie aus § 80 hervorgeht, die Discriminante der biquadratischen Gleichung bedeutet, deren Ausdruck durch die symmetrischen Grundformen wir hier nicht hinschreiben wollen.

Sodann adjungieren wir eine zu H gehörige Funktion. Eine solche ist z. B. $(x_0 - x_1)(x_2 - x_3)$, von der man leicht zeigt, dafs sie durch S_1, S_2 nicht geändert wird, durch T und T^2 und U neue Werte annimmt, speziell durch U den negativen Wert wie durch T. Allein es ist nicht nötig, dafs die zu adjungierende Funktion auch durch U seinen Wert ändert, weil U schon nicht mehr in der Gruppe G enthalten ist; wir brauchen nur eine Funktion zu wählen, die für die sämtlichen Permutationen aus G, die nicht zugleich H angehören, ihren Wert ändert. Statt daher die obige Funktion zu benutzen, können wir ebenso gut eine einfachere, nämlich

$$\psi(x_0, x_1, x_2, x_3) = x_0 x_3 + x_1 x_2,$$

wählen. Man erkennt, dafs

$$\psi_{S_1} = \psi_{S_2} = \psi_U = \psi,$$

dagegen

$$\psi_T = x_0 x_1 + x_2 x_3$$
$$\psi_{T^2} = x_0 x_2 + x_3 x_1$$

wird. Wollte man nun eine Funktion haben, die einer reinen kubischen Gleichung genügte, so müfste man, unter α eine primitive dritte Einheitswurzel verstanden, die Funktion

$$\psi + \alpha \psi_T + \alpha^2 \psi_{T^2}$$

betrachten. Für unsere Absicht hat dies aber keinen Zweck, im Gegenteil erschweren wir uns hierdurch nur die Übersicht, da die Rechnungen dann ziemlich umständlich werden, aufserdem aber nichts Neues bieten, da wir die kubischen Gleichungen ja prinzipiell schon erledigt haben. Daher halten wir an unserer Funktion ψ fest. Die kubische Gleichung für ψ hat die Form

$$(z - \psi)(z - \psi_T)(z - \psi_{T^2}) = 0$$

und läfst sich leicht bestimmen, denn es wird

$$\psi + \psi_T + \psi_{T^2} = f_2$$
$$\psi \cdot \psi_T + \psi \cdot \psi_{T^2} + \psi_T \cdot \psi_{T^2} = f_1 f_3 - 4 f_4$$
$$\psi \cdot \psi_T \cdot \psi_{T^2} = (f_1^2 - 4 f_2) f_4 + f_3^2,$$

wo f_1, f_2, f_3, f_4 die symmetrischen Grundformen von x_0, x_1, x_2, x_3 bedeuten. Somit hat die gesuchte Gleichung die Gestalt

$$z^3 - c_2 z^2 + (c_1 c_3 - 4 c_4) z - (c_1^2 c_4 - 4 c_2 c_4 + c_3^2) = 0.$$

Jetzt bilden wir eine zu K_1 gehörige Funktion, die sich also durch S_1 nicht ändert. Eine solche ist $x_0 + x_1$; diese geht durch S_2 in $x_2 + x_3$ über, was gleichfalls durch S_1 nicht geändert wird. Daher stellt die Differenz

$$\chi(x_0, x_1, x_2, x_3) = x_0 + x_1 - x_2 - x_3$$

eine Funktion dar, die zu K_1 gehört, durch S_2 aber das entgegengesetzte Vorzeichen annimmt und somit einer rein quadratischen Gleichung genügen mufs, wenn man φ und ψ adjungiert hat. Es wird nämlich

$$\chi^2 = f_1^2 - 4\psi - 4\psi_{T^2},$$

und es würde sich ψ_{T^2} rational durch φ und ψ ausdrücken

lassen. Beachten wir nun, dafs $TS_2 = S_1T$ ist, so ergiebt sich aus

$$\chi_{TS_2} = \chi_{S_1T} = \chi_T,$$

dafs χ_T eine zur Gruppe K_2 gehörige Funktion ist. Sie genügt der Gleichung

$$\chi_T^2 = f_1^2 - 4\,\psi_T - 4\,\psi.$$

Es erübrigt nun noch, x_0, x_1, x_2, x_3 durch die genannten Funktionen auszudrücken. Dabei ist es zweckmäfsig, auch noch die Funktion χ_{T^2} einzuführen, die sich rational durch die vorhergenannten ausdrücken läfst. Man findet sofort, dafs χ, χ_T, χ_{T^2} eine symmetrische Funktion ist, nämlich

$$\chi, \chi_T, \chi_{T^2} = c_1^3 - 4\,c_1\,c_2 + 8\,c_3,$$

und dann erhält man aus

$$\begin{aligned} f_1 &= x_0 + x_1 + x_2 + x_3 = c_1 \\ \chi &= x_0 + x_1 - x_2 - x_3 \\ \chi_T &= x_0 - x_1 + x_2 - x_3 \\ \chi_{T^2} &= x_0 - x_1 - x_2 + x_3 \end{aligned}$$

sofort in

$$\begin{aligned} x_0 &= \frac{1}{4}(c_1 + \chi + \chi_T + \chi_{T^2}) \\ x_1 &= \frac{1}{4}(c_1 + \chi - \chi_T - \chi_{T^2}) \\ x_2 &= \frac{1}{4}(c_1 - \chi + \chi_T - \chi_{T^2}) \\ x_3 &= \frac{1}{4}(c_1 - \chi - \chi_T + \chi_{T^2}) \end{aligned}$$

die Auflösung der biquadratischen Gleichung.